

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО

Факультет інформатики та обчислювальної техніки

(назва факультету, інституту)

Кафедра автоматизованих систем обробки інформації і управління

(назва кафедри)

"На правах рукопису"

УДК 519.876.5

«До захисту допущено»

В.о.завідувача кафедри

О.А.Павлов

(підпис)

(ініціали, прізвище)

“ ” 20 19 р.

МАГІСТЕРСЬКА ДИСЕРТАЦІЯ

на здобуття ступеня магістра

за спеціальністю 126 Інформаційні системи та технології

(код та назва спеціальності)

ОПП

Інформаційні управляючі системи та технології

(код та назва спеціалізації)

на тему: Моделювання системи консенсусу для децентралізованого контролю
якості контенту

Виконав: студент

VI курсу

ІС-81мп

(шифр групи)

Осипенко Олег Андрійович

(прізвище, ім'я, по батькові)

(підпис)

Науковий керівник

доц., к.т.н., доц. Жданова О.Г.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант

проф., д.т.н., проф. Томашевський В.М.

(науковий ступінь, вчене звання, прізвище, ініціали)

(підпис)

Рецензент

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській дисертації немає
запозичень з праць інших авторів без відповідних
посилань.

Студент

(підпис)

Київ – 2019

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ім. ІГОРЯ СІКОРСЬКОГО»**

Факультет інформатики та обчислювальної техніки
(повна назва)

Кафедра автоматизованих систем обробки інформації та управління
(повна назва)

Рівень вищої освіти другий (магістерський) за освітньо-професійною програмою

Спеціальність 126 Інформаційні системи та технології
(код і назва)

ОПП Інформаційні управляючі системи та технології
(код і назва)

ЗАТВЕРДЖУЮ
В.о.завідувача кафедри
О.А.Павлов
(підпис) (ініціали, прізвище)
«__» _____ 2019 р.

ЗАВДАННЯ

на магістерську дисертацію студенту
Осипенко Олегу Андрійовичу
(прізвище, ім'я, по батькові)

1. Тема дисертації Моделювання системи консенсусу для децентралізованого контролю якості контенту
- науковий керівник дисертації Жданова О. Г., доц., к.т.н., доц.
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
- затверджені наказом по університету від « 28 » жовтня 20 19 р. № 3770-с
2. Строк подання студентом дисертації « 2 » 12 20 19 р.
3. Об'єкт дослідження процес регулювання користувацького контенту в мережі Інтернет
4. Перелік завдань, які потрібно виконати виконати огляд існуючих підходів до регулювання користувацького контенту; розробити концепцію системи консенсусу; розробити модель системи консенсусу; запрограмувати модель системи консенсусу; експериментально підтвердити спроможність системи консенсусу; проаналізувати методи створення архітектури, яка забезпечить горизонтальне масштабування системи.
5. Орієнтовний перелік ілюстративного матеріалу динаміка зміни кількості

користувачів соціальних медіа; інформація про обсяги контенту, що генерується користувачами соціальних мереж; діаграма архітектури платформи системи; динаміка зміни обсягу грошей в системі при різних параметрах системи.

6. Орієнтовний перелік публікацій *стаття у фаховому журналі; тези доповіді в матеріалах науково-практичної конференції молодих вчених та студентів «Інформаційні системи та технології управління» (ІСТУ-2019).*

7. Консультанти розділів дисертації

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

8. Дата видачі завдання “ 2 ” вересня 20 19 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Строк виконання етапів магістерської дисертації	Примітка
1	<i>Систематизація результатів огляду літератури про регулювання контенту</i>	<i>20.09</i>	
2	<i>Постановка завдання дослідження</i>	<i>01.10</i>	
3	<i>Створення математичної моделі системи консенсусу</i>	<i>15.10</i>	
4	<i>Створення імітаційної моделі системи консенсусу</i>	<i>26.10</i>	
5	<i>Проведення та аналіз експериментів</i>	<i>01.11</i>	
7	<i>Опрацювання архітектури системи, що горизонтально масштабується</i>	<i>14.11</i>	
8	<i>Оформлення документації</i>	<i>17.11</i>	
9	<i>Подання роботи на попередній захист</i>	<i>20.11</i>	
10	<i>Подання роботи на основний захист</i>	<i>02.12</i>	

Студент

(підпис)

О. А. Осипенко

(ініціали, прізвище)

Науковий керівник

(підпис)

О. Г. Жданова

(ініціали, прізвище)

РЕФЕРАТ

Магістерська дисертація: 104 с., 37 рис., 22 табл., 32 джерела.

Актуальність. Ми живемо в епоху коли кожен, у кого є комп'ютер і підключення до Інтернету, може публікувати власні матеріали, доступ до яких можливий по всьому світу. Наряду з цим активний розвиток різноманітних соціальних мереж, що в першу є платформами для генерації користувацького контенту, створює нові виклики для регулювання віртуального простору. Згідно з дослідженнями The Global Change Data Lab, станом на 2018 рік спостерігається тенденція до постійного зростання кількості активних користувачів популярних соціальних мереж та платформ [дод. А, пл. 1]. Паралельно з ростом числа активних користувачів, збільшується об'єм контенту, що ними генерується. Згідно з інфографікою Influencer Marketing Hub, станом на 2019 рік кожні 24 години в соціальних медіа генерується: 4 000 000 000 постів в Facebook, 729 000 000 постів в Twitter, 78 000 000 фото в Instagram, 130 000 000 постів в Tumblr, 5 000 000 постів в Pinterest [дод. А, пл. 2]. Сьогодні ключовими гравцями у сфері регулювання користувацького контенту є уряди держав та компанії. Держави створюють правовий фундамент, на базі якого компанії формують правила, що регламентують процес користування їх продуктами. Однак, спроби застосування відмінних підходів у єдиному віртуальному просторі, створюють ряд проблем, які в кінцевому підсумку призводять до створення віртуальних кордонів та посягання на свободу слова в мережі Інтернет. У зв'язку з цим актуальною є розробка масштабованої системи, яка могла б забезпечити реалізацію єдиного та прозорого підходу для регулювання користувацького контенту.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась на кафедрі автоматизованих систем обробки інформації та управління Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» в рамках теми «Створення засобів імітаційного моделювання дискретно-подійних систем» (державний реєстраційний номер 0117U000923).

Мета дослідження - створення системи консенсусу для децентралізованого контролю якості контенту, яка забезпечить саморегулювання користувацького контенту в мережі Інтернет.

Для досягнення мети необхідно виконати наступні **завдання**:

- виконати огляд існуючих підходів до регулювання користувацького контенту;
- розробити концепцію системи консенсусу;
- розробити модель системи консенсусу;
- програмно реалізувати модель системи консенсусу;
- експериментально перевірити спроможність системи консенсусу;
- проаналізувати методи створення архітектури програмної платформи, яка забезпечить горизонтальне масштабування системи консенсусу, та спроектувати її.

Об'єкт дослідження - процес регулювання користувацького контенту в мережі Інтернет.

Предмет дослідження - методи та підходи до створення систем з децентралізованою архітектурою.

Наукова новизна отриманих результатів полягає у створенні нового підходу для систем, що повині приводити до кінцевого рішення із залучанням реальних користувачів. Запропонований підхід може створити підґрунтя для зміни методів адміністрування великих сучасних соціальних мереж, зміни підходів до державного управління та судової гілки влади.

Публікації. Статтю подано до друку у Вісник НТУУ «ХПІ», серія «Інформатика та моделювання». Тези опубліковано в матеріалах науково-практичної конференції молодих вчених та студентів «Інформаційні системи та технології управління» (ІСТУ-2019).

ДЕЦЕНТРАЛІЗОВАНІ СИСТЕМИ, КОНСЕНСУС ДЕЦЕНТРАЛІЗОВАНОЇ СИСТЕМИ, КОНТРОЛЬ ЯКОСТІ КОНТЕНТУ, МОДЕЛЮВАННЯ АЛГОРИТМУ КОНСЕНСУСУ

ABSTRACT

Master's thesis: 104 pp., 37 fig., 22 tabl., 32 sources.

Topicality. We live in an era in which everyone who has a computer and an internet connection can publish their own content, which is accessible worldwide. At the same time, the active development of various social networks, which primarily are platforms for user content generation, creates new challenges for virtual space regulation. According to research by The Global Change Data Lab, as at 2018 year there is a trend towards a steady increase in the number of active users of popular social networks and platforms. As the number of active users grows, the amount of content they generate is increasing. According to infographics by Influencer Marketing Hub, as at 2019 every 24 hours on social media is generated: 4 000 000 000 posts on Facebook, 729 000 000 posts on Twitter, 78 000 000 photos on Instagram, 130 000 000 posts on Tumblr, 5 000 000 posts on Pinterest. Today, key players in the field of user content regulation are governments and companies. States create a legal foundation, on the basis of which companies form the rules governing their product's using. However, attempts to use different approaches in a single virtual space create a number of problems that ultimately lead to emergence of virtual boundaries and encroachment on freedom of expression on the Internet. In that regard, it is important to develop a scalable system that could provide a unified and transparent approach to user content regulation.

Relation of work to scientific programs, plans, topics. The work was performed at the Department of automated information processing of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» within the topic «Creation of simulation tools for discrete event systems» (government registration number 0117U000923).

The purpose of the work - creation of consensus system for decentralized content quality control which will provide self-regulation of user-generated content on the Internet.

For achieving this goal, it is necessary to complete the following **tasks**:

- to consider existing approaches for user content regulation;

- to develop a concept of the consensus system;
- to develop a model of the consensus system;
- to implement a consensus system model;
- to validate experimentally the productivity of the consensus system;
- to analyze methods for creating an architecture of the platform which will allow horizontal scaling of the consensus system, and create the architecture.

The object of work - process of user content regulation on the Internet.

Subject of research - methods and approaches to creation of systems with decentralized architecture.

Scientific novelty of the obtained results is to create a new approach for systems, which should lead to the final decision with the involvement of real users. The proposed approach may provide a basis for changing the methods of administration of large modern social networks, changing approaches to public administration and the judiciary.

Publications The article is published in the Bulletin of NTUU "KPI", series "Informatics and Modeling". The abstracts are published in the materials of the scientific and practical conference of young scientists and students "Information systems and control technologies" (ISTU-2019).

DECENTRALIZED SYSTEMS, CONSENSUS OF THE DECENTRALIZED
SYSTEM; CONTENT QUALITY CONTROL, CONSENSUS ALGORITHM
MODELING

ЗМІСТ

1. СУЧАСНИЙ СТАН РЕГУЛЮВАННЯ КОНТЕНТУ В МЕРЕЖІ ІНТЕРНЕТ В УМОВАХ ГЛОБАЛІЗАЦІЇ.....	8
1.1 Огляд існуючих суб'єктів та механізмів оцінки контенту	8
1.2 Аналіз актуальних проблем оцінки контенту	17
1.3 Мета та завдання дослідження	26
1.4 Висновок.....	27
2.МОДЕЛІ ТА МЕТОДИ ДЛЯ СТВОРЕННЯ ДЕЦЕНТРАЛІЗОВАНОЇ СИСТЕМИ ОЦІНКИ ЯКОСТІ КОНТЕНТУ	29
2.1 Концепція системи консенсусу	29
2.2 Математична модель системи консенсусу	29
2.3 Розробка алгоритму розв'язання	30
2.4 Опис імітаційної моделі.....	32
2.5 Результати експериментальних досліджень	33
2.6 Висновок.....	38
3.АНАЛІЗ ТЕХНОЛОГІЙ ДЛЯ ЗАБЕЗПЕЧЕННЯ ГОРИЗОНТАЛЬНОГО МАСШТАБУВАННЯ	39
3.1 Паттерн Saga.....	39
3.2 Консиситентне хешування	46
3.3 База даних Apache Cassandra	57
3.4 Архітектура, що горизонтально масштабується	67
3.5 Висновок.....	69
4. РОЗРОБКА СТАРТАП-ПРОЕКТУ	70
4.1 Опис ідеї проекту	70
4.2 Технологічний аудит ідей проекту.....	74
4.3 Аналіз ринкових можливостей запуску стартап-проекту	75
4.4 Розробка ринкової стратегії проекту	87
4.5 Розробка маркетингової програми стартап-проекту	91
4.6 Висновок.....	93
ВИСНОВКИ.....	95
ПЕРЕЛІК ПОСИЛАНЬ	96
ДОДАТОК А Графічний матеріал	100
ПЛАКАТ 1 Динаміка зміни кількості активних користувачів соціальних медіа	101
ПЛАКАТ 2 Кількість контенту, що генерується користувачами соціальних медіа кожні 24 години та топ соціальних мереж по світу	102
ПЛАКАТ 3 Діаграма архітектури платформи системи консенсусу	103
ПЛАКАТ 4 Динаміка зміни об'єму грошей користувачів в часі при різній кількості суддів, 30% поганих користувачів та відсутності апеляцій	104

1 СУЧАСНИЙ СТАН РЕГУЛЮВАННЯ КОНТЕНТУ В МЕРЕЖІ ІНТЕРНЕТ В УМОВАХ ГЛОБАЛІЗАЦІЇ

1.1 Огляд існуючих суб'єктів та механізмів регулювання контенту

Закони та законопроекти про регулювання користувацького контенту обговорюються у багатьох юрисдикціях і часто викликають активну суспільну дискусію, оскільки державне регулювання може потенційно порушувати ті фундаментальні права людини, які особливо яскраво виражаються в соціальних мережах, зокрема, право на свободу слова, право на свободу зібрань, різні аспекти права на приватне життя тощо. Проте в таких юрисдикціях, як Німеччина, Великобританія, Китай і США, поступово з'являються масиви норм, які застосовуються до взаємодій користувачів в різноманітних інтернет-сервісах. На даний момент можна виділити дві основні моделі законодавчого регулювання.

Перша базується на створенні окремого консолідованого акта, об'єктом регулювання якого є всі сервіси, які підпадають під поняття «соціальна мережа»/«блог» тощо (яскравий приклад такого підходу – Network Enforcement Act в Німеччині, який вводить відповідальність соціальних мереж за невиконання вимог з видалення незаконного контенту, що надійшов від користувачів [1]);

Друга базується на застосуванні вже існуючого законодавства до соціальних мереж/блогів тощо і внесення в нього відповідних змін (наприклад, введення нових складових в блок кримінального законодавства). За цим підходом йде регулювання у великому числі юрисдикцій, що може бути пов'язано зі складністю визначення понять «соціальна мережа»/«блог» тощо - чи входять в нього комерційні сервіси типу Amazon? Так, наприклад, деякі види шахрайства або образи в соціальних мережах в Великобританії можуть потрапляти під дію Malicious Communications Act, в США Identity Theft Deterrence Act може регулювати випадки крадіжок паролів від акаунтів в соціальних мережах або інших сервісах.

Як бачимо, в деяких юрисдикціях вже прийняті заходи щодо забезпечення державного регулювання контенту, наряду з чим продовжуються обговорення можливого покращення існуючих механізмів регулювання та перспективи регулювання інтернет контенту в цілому. В Ірландії дискусія про необхідність вдосконалення кримінального законодавства до того стану, коли воно зможе застосовуватись до нейприйнятної поведінки в мережі Інтернет, ведеться приблизно з 2016 р. Одним із важливих результуючих документів цієї дискусії є доповідь Комісії по правовій реформі «Шкідливі комунікація та цифрова безпека» 2016 року [2]. Майже трьохсторінкова доповідь описує не лише можливі зміни законодавства щодо громадянської відповідальності за приведеній в документі список порушень, але і регламентує принципи, на яких має базуватись регулювання. При розгляді перспектив розвитку регулювання інформаційного простору необхідно в першу чергу мати на увазі, що введення обмежень на діяльність тих же соціальних мереж може порушувати деякі права і свободи людини, а також ставити під загрозу вільний обмін інформацією в світі. У цьому сенсі найяскравіший приклад регулювання соціальних мереж – німецький «Закон про соцмережі», який активно критикується навіть на міжнародному рівні в ОБСЄ та Раді Європи [3]. В нових технологічних умовах держави змушені перекладати свої обов'язки в сфері контролю та запобігання поширення забороненої і недостовірної інформації на самих суб'єктів, що поширюють (які організовують поширення) інформацію в мережі Інтернет, і самостійно здійснювати втручання в діяльність цих суб'єктів в мінімальній кількості випадків. Держави шукають нові підходи до поєднання державного і недержавного регулювання, а також нові способи примусу до виконання законодавства про поширення забороненої або недостовірної інформації.

Говорячи про застосування різних способів недержавного регулювання, необхідно відзначити різні види саморегулювання, якими є не лише регулювання через громадські об'єднання «за інтересами», а й інші види саморегулювання: наприклад, держава встановлює обов'язок для інформаційного посередника розробити політику протидії поширенню забороненої інформації або боротьби з

недостовірною інформацією і потім вимагає дотримання цієї політики під загрозою застосування санкцій. Говорячи про різні способи примусу до виконання законодавства про поширення забороненої чи недостовірної інформації, необхідно підкреслити, що в різних державах зустрічаються такі заходи впливу на порушників як штраф, блокування ресурсу, на якому розміщується неприйнятна інформація, заборона на розміщення реклами, уповільнення швидкості вхідного і/або вихідного трафіку тощо. Система видалення і блокування незаконного контенту в Німеччині заснована на саморегулюванні [4]. Саморегулювання в області незаконного контенту здійснюється в Німеччині на двох рівнях.

На першому рівні моніторинг і видалення/блокування незаконного контенту покладені на самі цифрові платформи (в тому числі соціальні мережі). Рішення про те, чи є контент незаконним, приймається самою платформою. Ефективність механізму саморегулювання з боку цифрових платформ забезпечується високими штрафами в разі, якщо платформа не реагує на інформацію, що надійшла зі скаргою про незаконний контент.

На другому рівні моніторинг незаконного контенту і прийняття рішення по тому, чи є контент законним, може здійснюватися уповноваженими СРО - саморегулювними організаціями. Підстави блокування сформульовані в загальному вигляді (наприклад, блокуванню підлягає інформація, поширення якої заборонено Кримінальним кодексом Німеччини; інформація, що розпалює ворожнечу за ознаками віросповідання тощо), в зв'язку з чим СРО та цифрові платформи самі повинні приймати рішення щодо того, чи вважають вони контент незаконним, керуючись при цьому прийнятими у відповідних СРО локальними актами. Іншими словами, цифрові платформи і СРО в Німеччині самі можуть визначати правила видалення контенту та його оцінки контенту з точки зору його законності. При цьому правом встановлювати критерії для оцінки контенту на предмет незаконності наділені і окремі органи державної влади.

З точки зору держав, ключове значення для дотримання прав людини, включаючи право на свободу слова і самовираження, грають міжнародні правові

документи. Одне з центральних місць серед них займає Міжнародний пакт про громадянські та політичні права [5]. Стаття 19 встановлює універсальні принципи свободи слова, актуальні і в цифрову епоху. Стаття 20(2) також встановлює обмеження свободи слова, пов'язані з проявами дискримінації та закликами до ворожнечі і насильства. На жаль, багато держав зловживають положеннями цієї статті, особливо при введенні антиекстремістського законодавства, що виявляється надзвичайно розмитим і створює можливості для численних зловживань.

Практично всі держави вимагають від компаній видаляти незаконний контент (дитячу порнографію, заклики до насильства і т.п.). Деякі держави навіть впроваджують повноцінні системи онлайн-цензури. Обмежувачий підхід до регуляції стає все більш поширеним, проявляючись в законах про екстремізм, наклепи, богохульство, фейкові новини і пропаганду. Також посилюється увага держав до регулювання конкретних онлайн-платформ. Традиційно багато держав приймали закони, які знімали з компаній відповідальність за контент, який публікується третіми сторонами на їх платформах. Однак сьогодні багато держав накладають на компанії зобов'язання по обмеженню користувацького контенту з можливим застосуванням санкцій до компаній-порушників (зазвичай це великі штрафи для юридичних осіб). Збільшилася також кількість країн, які зобов'язують компанії здійснювати моніторинг і фільтрацію користувацького контенту, що підриває основи прав людини. Одним з найбільш яскравих прикладів є прийняття в Німеччині закону "про охорону прав в соцмережах" (NetzDG) [6]. Відповідно до цього закону соціальні медіа повинні протягом доби після отримання запиту видаляти заборонений контент. У разі зволікання на них можуть бути накладені штрафи, що досягають 50 млн. євро. Незважаючи на те, що побоювання держав щодо питань конфіденційності та національної безпеки зрозумілі, прийняті ними регуляторні заходи часто виявляються надлишковими і несуть в собі непередбачені ризики. Одним з найбільш значущих ризиків є делегування функцій держави приватним компаніям, які керуються виключно економічними стимулами і володіють дуже слабкою підзвітністю інститутам

громадянського суспільства. Деякі держави вимагають здійснювати видалення "небажаного" контенту на екстериторіальній основі. Це може привести до суттєвих обмежень свободи слова, що не прив'язані до національних кордонів, і так званим «суверенізованим» сегментам інтернету. Очевидно, що в рамках існуючих міжнародних правових стандартів, запити на видалення контенту повинні розглядатися окремо в кожній юрисдикції, не зачіпаючи іноземних громадян і не порушуючи зв'язності інтернету. Варто також відзначити, що запити більшості держав часто і зовсім не ґрунтуються на національному законодавстві. Замість ініціювання судової процедури в рамках власної національної юрисдикції держави вимагають видалення контенту на основі корпоративних умов надання послуг користувачам онлайн-платформи. Компанії охоче задовольняють ці запити, побоюючись можливих втрат прибутку. Лише деякі держави створили спеціальні органи і виробили чіткі нормативні процедури для примусу компаній до видалення незаконного контенту. Але і в цих випадках ІТ-компанії тісно співпрацюють навіть з найавторитарнішими режимами.

Щодо компаній, сьогодні вони є основними операторами платформ для комунікації та обміну думками, отримання доступу до інформації, торгівлі та розвитку людського потенціалу. Компанії збирають персональні дані у мільярдів жителів Землі, багато хто з яких лише ситуативно користуються інтернетом. Такі дані охоплюють широке коло аспектів життя людини: від споживчих переваг до громадянської активності. Більшість компаній офіційно декларують відданість цінностям сталого розвитку і вільного суспільства. Проте насправді їх політика щодо конфіденційності користувачів і надання доступу до інформації викликає численні запитання. Схоже, що зазвичай компанії готові дотримуватися прав людини тільки в тій мірі, в якій їх зобов'язує це робити національне законодавство. Величезного значення в цьому контексті набуває відсутність будь-якого універсального міжнародного правового документу, який міг би встановити загальні вимоги ведення бізнесу для ІТ-компаній, сумісні з дотриманням основних прав. Хоча в 2011 році ООН прийняла Керівні принципи підприємницької діяльності в аспекті прав людини, рекомендаційний характер

даного документу дозволяє компаніям просто ігнорувати його [7]. Проте, він передбачає набір важливих принципів, які можуть бути дуже актуальні для ІТ-індустрії. Згідно з даним документом компанії повинні як мінімум робити наступне:

- а) уникати прямого порушення прав людини або сприяння таким порушенням;
- б) прийняти зобов'язання на високому рівні щодо дотримання прав користувачів сервісів компанії;
- в) проводити оцінку потенційного негативного впливу діяльності компанії на дотримання прав людини;
- г) виробити стратегії пом'якшення можливих негативних наслідків діяльності компанії для дотримання прав людини при роботі в юрисдикціях з суперечливими вимогами місцевого законодавства;
- д) проводити регулярний аналіз зусиль компанії щодо дотримання прав людини шляхом регулярних консультацій із зацікавленими сторонами і широкою громадськістю;
- е) забезпечити функціонування належних механізмів захисту прав користувачів (в тому числі механізмів оперативної подачі скарг) для того, щоб вони не відчували себе безпорадними.

Кожна компанія намагається відповідати вимогам законодавства країни, в якій вона веде бізнес, що можна пояснити природним прагненням до максимізації прибутку та підтримання належного іміджу. Однак, в тих випадках, коли національне законодавство систематично порушує права людини, компанії лише погіршують становище речей, намагаючись виконувати вимоги регуляторів. Наприклад, часта смислова невизначеність антиекстремістських законів дозволяє державам вимагати від компаній видалення самих різних типів контенту для продовження нормальної роботи в країні. Міжнародна некомерційна асоціація ІТ-компаній The Global Network Initiative розробила додаткові рекомендації по використанню інструментів мінімізації негативного впливу репресивного національного законодавства на кінцевих користувачів [6].

Одним з таких інструментів є максимізація прозорості. Компанії виробили практику публікації звітів про прозорість, які містять агреговані дані про державні запити на видалення контенту і користувацьких даних. До теперішнього моменту багато компаній щорічно повідомляють про кількість державних запитів, які вони отримують і задовольняють. Таким чином компанії демонструють ті види тиску, з якими вони стикаються при модерації контентного наповнення. Звіти дозволяють оцінити число запитів на видалення контенту в рамках офіційних правових процедур, число випадків задоволення таких запитів і опис правових норм, якими обґрунтовуються такі рішення. Однак компанії вкрай нерегулярно розкривають детальну інформацію про те, як вони реагують на такі запити. Крім того, компанії практично не розкривають інформації про внутрішні корпоративні стандарти політики щодо контенту. На даний момент компанії майже не публікують відповідну звітність, хоча, очевидно, що внутрішні правила обумовлюють переважну більшість випадків видалення контенту.

Експоненціальне зростання контенту, що створюється користувачами, призводить до істотного ускладнення умов надання послуг на онлайн-платформах і виникнення спеціальних структур всередині ІТ-компаній, що займаються питаннями модерації контенту. В даний час в розробку корпоративної політики щодо контенту залучаються не тільки юристи і фахівці в області публічної політики, а й керівники вищої ланки. Компанії створюють спеціальні групи «довіри і безпеки» (займаються питаннями зловживання на онлайн-платформах) і навіть антитерористичні відділи, які тісно співпрацюють з національними правоохоронцями. Однак при такому зростанні інфраструктури регулювання контенту, корпоративні стандарти щодо заборонених матеріалів на онлайн-платформах залишаються дуже розпливчастими. Це стосується не тільки протидії тероризму і масовим актам насильства, а й таких питань як прояв ненависті, домагання, зловмисне використання онлайн-платформ тощо. Стандарти протидії різним зловживанням, прийняті на великих онлайн-платформах, часто декларативні і не мають під собою чітко визначених формальних підстав. Незважаючи на зусилля компаній з протидії так званих

«мові ненависті» і придушенню меншин, такі акти продовжують відбуватися з високою регулярністю. Якщо компанії продовжать поточну політику по боротьбі з такими негативними проявами, свобода слова може бути жорстко обмежена навіть у тих випадках, коли вони будуть неочевидні або відсутні зовсім. Щоб уникнути таких ризиків, компаніям слід чітко визначити формальні підстави для обмежень, що вводяться і продемонструвати необхідність і доречність дій, що застосовуються у відношенні користувацького контенту.

На даний момент компанії використовують наступні інструменти модерування контенту:

- а) автоматичні системи модерації (автоматичне маркування контенту, видалення і попередня фільтрація, обробка природно-мовних текстів тощо);
- б) маркування контенту кінцевими користувачами (спільнотою користувачів онлайн-платформи і/або спеціально навченими модераторами з числа користувачів);
- в) маркування контенту співробітниками компанії;
- г) деактивація акаунтів порушників;
- д) розсилка повідомлень про порушення правил або видалення користувацького контенту (часто компанії або зовсім не розсилають повідомлення користувачам, які розмістили такий контент, або лише інформують про факт видалення, не залишаючи можливості для оскарження цього рішення).

Багато з вищезазначених інструментів спочатку застосовувалися головним чином для захисту правовласників, однак згодом область їх застосування була істотно розширена, що дозволило створити масштабні системи попередньої модерації контенту і цензури. Національні закони часто не підходять для компаній, які потребують загальних норм для своїх географічно і культурно різноманітних платформ, що в першу чергу націлені на задоволення інформаційних потреб користувача. Але права людини, якщо вони реалізуються прозоро і послідовно, зі значним внеском самих користувачів і громадянського суспільства, забезпечують основу для того, щоб держави і компанії були

підзвітні користувачам незалежно від державних кордонів. Принципи прав людини можуть дозволити компаніям створити інклюзивне середовище, яке враховує різноманітні потреби та інтереси своїх користувачів.

Держави прагнуть сформувати правове середовище, в рамках якого компанії можуть проводити модерацію контенту, в той час як компанії засновують доступ до своїх платформ на базі визначеної для користувача угоди з умовами, що встановлюють стандарти допустимого вираження своєї думки на цих платформах. Виходячи з наявного вже зараз регуляторного досвіду і обговорення майбутніх моделей регулювання, можна виділити такі релевантні питання:

а) за якою моделлю має здійснюватись регулювання мережевого контенту, і чи релевантно взагалі створювати законодавчі акти, які не зможуть охопити всі аспекти використання соціальних сервісів?

б) як регуляторам дотримуватись балансу між публічним інтересом, що вимагає державного втручання в роботу соціальних мереж, і правами користувачів в Інтернеті?

в) наскільки користувацький контент взагалі повинен регулюватися державою і яка роль саморегулювання?

Ці питання відображають класичну проблему, яка все частіше постає перед правом в епоху нової цифрової економіки - право не встигає за стрімким розвитком технологій, через що утворюється все більше і більше «дірок» в регулюванні. У випадку з соціальними мережами це ускладнюється глобальним охопленням таких платформ як Facebook і Instagram, де інформація поширюється на величезну аудиторію за вкрай малий проміжок час. У зв'язку з цим регулювання не повинно обмежуватися правовими інструментами - їх необхідно доповнювати співпрацею з самою індустрією, а також довгостроковими заходами, спрямованими на підвищення медіаграмотності та соціальної відповідальності власників соціальних мереж і їх користувачів [8]. Враховуючи всі вищеописані аспекти, регулювання через самих користувачів є гнучким та ефективним рішенням, яке могло б врівноважити впливи законодавчих баз різних країн на єдиний контентний простір та вирішити похідні від цього явища

проблеми, які доводиться вирішувати провайдерам платформ, на яких здійснюється інформаційний обмін.

1.2 Аналіз актуальних проблем регулювання контенту

Ми живемо в епоху, коли кожен може брати участь в «діалозі» більш активно, ніж будь-коли. Сьогодні свобода ЗМІ і свобода слова означають надання можливості спілкуватися і ділитися інформацією всім і кожному. Кожен, у кого є комп'ютер і підключення до Інтернету, може публікувати власні матеріали, доступ до яких можливий по всьому світу. Паралельно з розвитком можливостей для наповнення інформаційного простору розвиваються і ризики, пов'язані з використанням цього простору. І чим більше розширюється інформаційний простір, тим складніше (а то і зовсім неможливо) здійснювати його регулювання. До настання епохи інтернету, інформацію і контент можна було головним чином отримати за допомогою мовних і друкованих засобів масової інформації, в зв'язку з чим вона легко контролювалася національними органами влади. Сьогодні веб 2.0 глобалізував інформацію. Таким чином, національні уряди втратили частину контролю над контентом, який споживають і продукують громадяни. При тому, що Інтернет різко розширив перспективи свободи вираження думки, висока доступність контенту різного характеру, включаючи контент порнографічного характеру або контент, шкідливий для дітей, викликає серйозну заклопотаність. Багато хто заявляє, що інтернет-контент вимагає більш суворого регулювання, що є практично неможливим, оскільки Інтернет є глобальним, відкритим середовищем, і немає єдиного розуміння правил, які повинні були б регулювати його на міжнародному рівні.

Широка доступність потенційно шкідливого і незаконного контенту в Інтернеті призвела до зростання стурбованості серед органів влади і суспільства. Навіть якщо переваги від вільної передачі інформації переважають недоліки, відповідальність за незаконний і шкідливий контент, який можна в ньому знайти, залишається серйозним питанням, особливо враховуючи той факт, що Інтернет дозволяє зберігати анонімність. Як і у випадках з багатьма іншими продуктами і послугами, зловживання/використання інформаційного простору не за цільовим

призначенням може містити потенційну небезпеку. Потенційні ризики, з якими може зустрітися користувач інформаційного простору:

а) нелегальний контент, наприклад, матеріали, що демонструють жорстокість і насилля; матеріали, що містять інструкції для здійснення злочинів або вчинення протизаконних дій тощо;

б) контент, що не призначений для перегляду особами, які не досягли 18 років - наприклад, порнографія, матеріали сексуального характеру або інший контент для дорослих, який може бути недоречний для перегляду дітьми і підлітками;

в) контент, спрямований на пропаганду війни і розпалювання національної, расової чи релігійної ненависті; контент, спрямований на заперечення, надзвичайну мінімізацію, схвалення або виправдання;

г) геноциду або злочинів проти людства, сюди ж відноситься контент із закликами до сепаратизму, публічним виправданням тероризму, пропагандою наркотиків, публікаціями про способи самогубства, закликами до здійснення насильницьких дій проти конкретних людей або групи людей тощо;

д) контент, що містить дані, що посягають на честь, гідність чи репутацію іншої особи, а також містять будь-яку недостовірну інформацію; сюди ж відноситься контент, що містить заклики до образ та погроз по відношенню до конкретних користувачів.

Процес регулювання користувацького контенту на сьогоднішній день представлений двома сторонами: зовнішня і внутрішня цензура. З одного боку, зовнішня цензура (держави, корпорації тощо) керуючись певними загальноприйнятими уявленнями про так званий «шкідливий» контент здійснює регулювання мережевого контенту, що генерується користувачами. Держава як головний суб'єкт законотворення створює правовий фундамент, на базі якого корпорації формують правила, що регламентують процес користування їх сервісами/продуктами. З іншого боку, внутрішня цензура самих інтернет-користувачів, яка обмежується індивідуальними уявленнями про «прийнятне/неприйнятне» напряду здійснює регулювання мережевого контенту ще на етапі його створення. Так, в більшості випадків, якщо користувач

соціальної мережі розуміє, що публікація певного контенту може спровокувати скарги, він відмовляється від ідеї публікувати цей контент. А у випадку появи неприйнятного за думкою більшості користувачів контенту вони створюють скарги, які допомагають виявити такий контент і прийняти відповідних заходів. Тобто, основна частина роботи по фільтрації контенту здійснюється самими користувачами. Однак, фінальний і найважливіший етап (прийняття рішення про задоволення або відхилення скарги) лягає на плечі одного робітника служби підтримки. Виходячи з вищеописаного, існують такі проблеми.

Незалучення користувачів до процесу прийняття рішення - остаточне рішення про неприйнятність того, що більшістю оцінюється як неприйнятне, приймається однією особою. Таким чином, проблема, що виникла в результаті агрегації індивідуальних поглядів колективу, вирішується конкретним індивідумом з власними індивідуальними поглядами. При такому підході усереднене твердження, яке представляє думку більшості замінюється на випадкове твердження, що представляє думку конкретної особи.

Суб'єктивізм - аналогічно тому, як кожна людина має індивідуальне світосприйняття і норми моралі, у кожної людини є власні уявлення про сутність радикальних політичних матеріалів; матеріалів, що виражають ненависть або закликають до насилля; матеріалів, що трансюють расистські та ксенофобні погляди тощо.

Вплив людського фактору - щоденно працівник служби підтримки розглядає сотні скарг різного характеру. Очевидно, що на різних етапах роботи уважність, ясність та логічність мисленнєвих процесів і якість роботи в цілому різняться. Ця особливість створює додаткові можливості для впливу випадкових факторів на результат розгляду скарги.

Відсутність єдиного розуміння правил регулювання мережевого контенту - подібно законодавству різних держав, неможливо сформувати єдиний підхід до регулювання інтернет простору, який буде прийматись на міжнародному рівні. Принаймні, на сьогоднішній день. Тому регулювання через самих користувачів інтернет-простору представляється можливим вирішенням

проблеми. По-перше, такий підхід дозволяє транслювати думку більшості. По-друге, такий механізм є більш гнучким, ніж державне регулювання. По-третє, саморегулювання - єдиний механізм, який сприяє свободі слова. Тут важливо розуміти, що право на свободу вираження думки застосовується в однаковій мірі до користувачів і громадян як в інтернеті, так і за його межами (стаття 20 резолюції ООН про «Заохочення і захист всіх прав людини, громадянських, політичних, економічних, соціальних і культурних прав, включаючи право на розвиток»).

Відмінність законодавчого апарату різних країн - незважаючи на базові поняття щодо законності/незаконності конкретних випадків, закони різних країн можуть дуже сильно відрізнятись між собою. Відмінність законодавчої бази тягне за собою відмінність уявлень про норми поведінки в суспільному та віртуальному середовищі. Як наслідок, різні соціальні сервіси можуть мати кардинально різні правила щодо розміщення матеріалів того чи іншого характеру. Наряду з віртуальною глобалізацією це створює велику проблему. Наявність індивідуальних для кожного сервісу/платформи положень щодо публікації контенту не заважає урядам інших країн створювати запити на видалення неприємного з їх точки зору контенту з конкретного сервісу. Так, компанія Google у 2012 році опублікувала в мережі «Звіт про прозорість» [9]. У цьому звіті описані деякі ситуації, пов'язані із запитами певних країн провилучення контенту, і описано, як компанія Google реагувала на такі запити. У зв'язку з відомим роликком до фільму «Невинність мусульман» компанія Google 15 вересня 2012 року заявила: «Ми докладемо зусиль для того, щоб створити таке співтовариство, в якому комфортно всім і яке також дозволяє людям висловлювати різні думки. Це може викликати великі труднощі, оскільки те, що в одній країні вважається нормальним, в іншій може бути визнано образливим. Цей відеоматеріал - який легко знайти в мережі - чітко вписується в рамки наших вимог, тому ми не будемо видаляти його з YouTube. Однак ми обмежили доступ до нього в тих країнах, в яких він вважається незаконним, наприклад, в Індії та Індонезії, а також в Лівії і в Єгипті [9]. Крім цього випадку, компанія Google приділяє багато уваги розгляду запитів на видалення контенту від різних

державних установ. Уряди звертаються до Google із запитом на видалення вмісту з різних причин. Деякі запити містять звинувачення в наклепі, в інших вказується, що вміст порушує місцеве законодавство, яке забороняє ворожі висловлювання чи вміст для дорослих. Закони пов'язані з цими питаннями, відрізняються залежно від країни. Команда Google відносно оцінює кожен запит і перевіряє публікацію, щоб визначити, чи потрібно її видалити через порушення місцевого законодавства або політики Google щодо контенту [10]. Вищенаведені факти свідчать про те, що деякі компанії готові йти назустріч уряду конкретних держав. Однак, в їх випадку це збільшує кількість ресурсів, що витрачається на регулювання контенту і в певній мірі обмежує контентне наповнення ресурсу законодавством інших країн (адже у випадку підтвердження скарги публікація видаляється). Публікація може і приховуватись від громадян конкретних країн, проте це при масштабуванні створює ряд додаткових робіт, що має виконуватись службою підтримки. Таким чином, врахування законодавства інших країн при регулюванні контенту можуть собі дозволити лише компанії, які турбуються про свій імідж на міжнародному рівні і мають на це достатньо ресурсів. Крім того, Інтернет сьогодні є найважливішою формою здійснення права на свободу вираження поглядів та права на свободу інформації. Тому будь-які національні закони не повинні негативно позначатись на здатності здійснювати перелічені права в мережі громадянами інших країн.

Державний контроль - ця проблема є продовженням вищеописаної. Якими б не були положення щодо регулювання інтернет-простору в межах певного сервісу, всі вони спираються на закон тієї країни, за законодавством якої даний сервіс був створений. В тих випадках, коли національне законодавство систематично порушує права людини, компанії лише погіршують становище речей, намагаючись виконувати вимоги регуляторів. Наприклад, часта смислова невизначеність антиекстремістських законів дозволяє державам вимагати від компаній видалення найрізноманітніших типів контенту. Таким чином, прив'язка до національних законів щодо свободи вираження думок в мережі порушує єдність інформаційного простору і сприяє створенню віртуальних

кордонів, що суперечить фундаментальній характеристиці мережі Інтернет - відкритість. Саме ця характеристика робить Інтернет таким популярним середовищем. Саме ця відкритість дає Інтернету його стійкість і представляє собою цінність для громадянських свобод і фактично - цінність для економіки. Інтернет являє собою мережу взаємопов'язаних мереж - відкриту архітектуру, яка пропонує інфраструктуру комунікації «будь-кого з будь-ким». Саме тому надзвичайно важливо підтримувати цю відкритість і розвивати механізми саморегулювання, які дозволили б мінімізувати втручання держав, але при цьому ефективно регулювати контентний простір. Ефективність при регулюванні контенту підвищується при наявності просторів з єдиними правилами, тому важливі заходи з вибудовування загальних підходів до регулювання контенту.

Відсутність прозорості у прийнятті рішень - користувачі інформаційного простору не знають, як приймаються рішення про задоволення або відхилення їх скарг. З іншого боку, користувачі не убезпечені від того, що прийнятий для них контент не буде видалений через неприйнятність цього контенту для уряду інших держав.

Виправданість регулювання віртуального простору на державному рівні - якщо порівняти швидкість розвитку технологій та створення законів, то виникають сумніви щодо необхідності регулювання інтернет-середовища на законодавчому рівні. Державні органи багатьох країн намагаються реагувати на існування Інтернету за допомогою звичних для державного апарату засобів, наприклад шляхом прийняття парламентами відповідних законів або винесення судами рішень у справах про доступ до протизаконного контенту. У більшості випадків такі спроби бажаних результатів не приносять. Ефект застосування тієї ж цензури щодо неприйнятної інтернет контенту обмежений, оскільки динамічність та мінливість Інтернет простору часто призводить до того, що програми фільтрації легко можна обійти. Фактично це доводить слабкість застосування традиційного підходу законотворення в тих ситуаціях, коли держава намагається регламентувати Інтернет.

На сьогоднішній день більшість країн взагалі не мають чіткої законодавчої бази щодо регулювання інтернет простору. Щодо решти країн, розвиток вже

існуючих нормативно-правових баз щодо запобігання негативному впливу інформаційного простору ускладнюється через надвисоку швидкість соціально-економічних змін, викликаних використанням нових технологій, появою нових видів економічної діяльності та розвитком суспільства в цілому. Як наслідок, цифровізація вимагає від законодавчих органів все більш високої швидкості прийняття рішень і глибинного розуміння трендів в соціально-економічному і технологічному розвитку, що є практично неможливим. Мова йде про постійну адаптацію законів, які в кінцевому підсумку все одно будуть застарілими порівняно з віртуальним середовищем, що постійно змінюється. Підтримання статичних за своєю природою законів для такого динамічного середовища як Інтернет є складним викликом. В даному випадку залучення представників цього динамічного середовища в якості суб'єктів його регулювання є вдалим рішенням.

У сфері контент-регулювання досі не знайдені універсальні рецепти для вирішення наступних питань:

а) до яких меж повинна зберігатися колишня роль держави як суб'єкта, який встановлює правила і забезпечує їх примусове застосування в рамках контент-регулювання (особливо з урахуванням транскордонного характеру обміну інформацією та знаходження інформаційних посередників за межами юрисдикції держав, що встановлюють конкретні правила гри);

б) які з існуючих способів примусового впливу на суб'єктів з метою спонукання до виконання вимог держави (виконання інформаційного законодавства) працюють і які нові способи можуть бути ефективно використані в рамках нових технологій;

в) в якій мірі можна залучати до регулювання (встановлення правил наповнення інформаційного простору контентом та контролю за їх дотриманням) безпосередніх суб'єктів інформаційних відносин (у тому числі споживачів інформації) ;

г) які економічні і соціальні стимули (як позитивні, так і негативні) можуть бути використані для спонукання суб'єктів інформаційних відносин до дотримання правил регулювання контенту.

Впровадження Інтернету відбувалося одночасно з впровадженням правил і моделей поведінки, які виробляли самі користувачі. Саме це відрізняє Інтернет від всіх існуючих раніше технологій зв'язку. З достатньою часткою обґрунтованості такі процеси можна вважати процесами практичного саморегулювання. «Нетікет» (мережевий етикет) став першим неформальним кодексом поведінки, який був вироблений не представниками відповідної галузі, а самими користувачами, що прагнули виористовувати мережу в цивілізованих рамках. Таку ідеологію слід розвивати і популяризувати серед користувачів мережі. Вона ж повинна лежати в основі інших форм регулювання Інтернету. Цілком очевидно, що користувачі, які задають вектор розвитку середовища, можуть забезпечувати регулювання цього середовища. Процес регулювання Інтернету породжує необхідність нових і навіть новаторських підходів. Традиційний процес законотворення, головними суб'єктами якого є держава, її законодавчі органи, бюрократичний апарат і судова система, в більшості випадків вже не відповідає сучасним вимогам. Це пояснюється двома обставинами: по-перше, регулювати Інтернету законами у багатьох випадках не представляється можливим і, по-друге, Інтернет як засіб глобальної комунікації не вписується в рамки національних кордонів.

Питання про те, хто повинен регулювати користувацький контент, активно обговорюється. Загальна згода щодо прав та обов'язків людей в розрізі користування інтернетом поки що не досягнута. Однак, той факт, що сатиричні відеоролики/пости, розміщені в Інтернеті, отримують широкий розголос, демонструє поступовий відхід від звичайного підходу до регулювання користувацького контенту і піднімає питання корпоративної соціальної відповідальності. Мінімізація впливу державного контролю на процес регулювання інтернет середовища має вітатися, оскільки це дозволяє більшій кількості людей вільно виражати свої погляди. З іншого боку, коли справа доходить до саморегулювання, до користувачів пред'являються більш високі

вимоги, адже вони мають турбуватись про те, як той чи інший контент буде сприйнятий певною вибіркою людей. Це сприятливо впливає на процес фільтрації контенту. По-перше, при такому підході «людський суд» є «фільтром второго порядку», який здатний відфільтровувати все те, що пройшло «фільтр першого порядку» (контент, який не порушує очевидні, базові правила щодо характеру контенту, що публікується), але при цьому є шкідливим/образливим. Наприклад, текстовий пост, що завуальовано пропагандує жорстокість та нетерпимість або принижує честь та достоїнство певної особи. По-друге, користувач, який плануватиме опублікувати той чи інший матеріал, буде розуміти, що потенційно спірний контент з більшою вірогідністю буде оскаржений комуніою, яка окрім закону транслює ще й узагальнені норми моралі. Таким чином, генерований користувачами контент буде більш свідомим і саморегулювання підвищить ступінь фільтрації контенту ще на етапі його створення. Спільна робота зацікавлених сторін (самих користувачів, які беруть участь в оцінці контенту) в поєднанні з усвідомленням та виконанням прийнятих на себе обов'язків (вдумлива публікація контенту) в кінцевому підсумку можуть забезпечити значно безпечніший, але в той же час гнучкіший інформаційний простір. Вищеописаний підхід пропагандує спільну роботу зацікавлених осіб як найбільш ефективний спосіб управління потенційними ризиками інформаційного простору.

У власному розумінні цього слова «регулювання» передбачає досить довільний процес, що здійснюється під керівництвом держави, функції якої зазвичай зосереджуються в межах якого-небудь (більш-менш) незалежного регламентарного органу, який приймає рішення в разі конфлікту інтересів. Процес прийняття рішень настільки складний, що певний спеціалізований орган, що складається з незалежних експертів, впорається з цим краще, ніж державні чиновники. Подібно вирішенню спірних ситуацій в реальному житті, спірний контент може оцінюватись судом присяжних. Однак не реальним, а віртуальним (враховуючи той факт, що інцидент стався в віртуальному середовищі). Наприклад, коли користувач за допомогою конкретного сайту поширює

неприйнятний контент, адміністрація сайту не має права приймати на себе функції правоохоронних або судових органів і об'єктивно не в змозі оцінити, чи є той чи інший контент, розміщений на сайті, законним. В даному випадку адміністрація може оцінити контент лише з точки зору бачення компанії керуючись власним набором положень. Такий підхід є локальним і не має єдиного механізму регулювання. Застосування віртуального суду присяжних натомість забезпечить глобальний, єдиний механізм врегулювання спірних питань щодо контенту інтернет простору і мінімізує поділ глобальної мережі Інтернет на національні сегменти. Розвиток нових форм міжнародної співпраці, саморегулювання та корпоративної відповідальності є надзвичайно важливим. Адже Інтернет як глобальне середовище представляє собою важливий інструмент пошуку, отримання та розповсюдження інформації, що не залежить від кордонів. Тому важливо, щоб це середовище залишалось єдиним та вільним.

1.3 Мета та завдання дослідження

Метою даної роботи є створення системи консенсусу для оцінки якості контенту, яка забезпечить саморегулювання користувацького контенту в мережі Інтернет.

Для досягнення мети необхідно вирішити наступні завдання:

а) розробити концепцію системи консенсусу:

- 1) визначити основні компоненти системи та специфіку їх роботи;
- 2) розробити механізм збору скарг;
- 3) розробити механізм визначення суду присяжних;
- 4) розробити механізм прийняття рішень судом присяжних;
- 5) розробити механізми стягнення штрафів та надання винагород.

б) програмно реалізувати модель системи консенсусу;

в) експериментально перевірити спроможність системи консенсусу;

г) проаналізувати методи створення архітектури програмної платформи, яка забезпечить горизонтальне масштабування системи консенсусу, та спроектувати її.

1.4 Висновок

В розділі розглянуто основні суб'єкти та підходи у сфері державного регулювання користувацького контенту та наведено приклади законодавчого регулювання контенту в різних країнах. Також розглянуто основні суб'єкти та підходи у сфері недержавного регулювання користувацького контенту та наведено приклади здійснення саморегулювання контенту в різних країнах. Описано роль компаній в регулюванні користувацького контенту та інструменти, які вони використовують для його регулювання. Розглянуто взаємодію між наведеними суб'єктами регулювання контенту, зокрема описано вплив законодавчих органів на процеси регулювання контенту у компаніях.

Виходячи з розглянутих механізмів та підходів, сформульовано перелік проблем, що є актуальними для сфери регулювання користувацького контенту, зокрема: суб'єктивізм та вплив людського фактору у процесі прийняття рішень; відсутність прозорості та незалучення користувачів до процесу прийняття рішень; відмінність законів та підходів до регулювання контенту; відсутність єдиного розуміння процесу регулювання контенту; надмірне втручання деяких держав у процес регулювання контенту в глобальному Інтернет просторі.

Виходячи з описаних проблем, оптимальним рішенням обрано створення технології, яка могла б забезпечити саморегулювання користувацького контенту. В розділі описано концепцію віртуального суду присяжних, що здатна забезпечити глобальний та неподільний механізм врегульовування спірних питань щодо контенту і мінімізувати поділ глобальної мережі Інтернет на національні сегменти.

В розділі наведено мету, що полягає у створенні системи консенсусу для оцінки якості контенту та завдання дослідження, зокрема: розробка концепції та моделі системи консенсусу; визначення значень параметрів системи, що забезпечать її ефективну роботу; аналіз методів створення архітектури системи, що забезпечить її горизонтальне масштабування.

2 МОДЕЛЬ ДЕЦЕНТРАЛІЗОВАНОЇ СИСТЕМИ ОЦІНКИ ЯКОСТІ КОНТЕНТУ ТА ЇЇ АНАЛІЗ

2.1 Концепція системи консенсусу

Необхідно розробити алгоритм консенсусу, який би надавав вирішення проблеми судочинства щодо скарги на контент. Алгоритм повинен бути кінцевим (таким, що виконується за скінченний час) та дискретним у результаті (скарга визнається коректною та автор некоректного посту штрафується, а особа, що поскаржилася, та всі учасники судочинства отримують винагороду, що взята зі штрафної суми). Або ж скарга визнається некоректною, і людина що поскаржилася штрафується за неправомірну скаргу, а всі учасники судочинства отримують винагороду. Окрім цього, існує сенс передбачити систему апеляцій. Апеляцію може подати учасник судочинства, запустивши новий раунд пошуку рішень щодо скарги, але вже із залученням більшої кількості людей. Якщо апеляція признається неправомірною, штрафується людина, яка була ініціатором апеляції. Якщо ж апеляція визнана правомірною, штраф за попереднє рішення інвертується щодо учасників. У базовій версії системи, яка розробляється в межах магістерської дисертації, можна подати лише одну апеляцію.

2.2 Математична модель системи консенсусу

Нехай U - множина всіх користувачів системи, $N_C \in U$ - користувач, який створив скаргу, $N_R \in U$ - користувач, на якого створено скаргу. Для кожного етапу судочинства будемо мати множини:

- а) $E = \{N_1, N_2, \dots, N_k\}$ - множина випадково обраних користувачів-суддів для цього етапу (k - непарне число);
- б) $A \subseteq E$ - підмножина множини E , яка містить користувачів, які підтримали скаргу;
- в) $R \subseteq E$ - підмножина множини E , яка містить користувачів, які не підтримали скаргу.

Якщо $|A| > |R|$ (кількість користувачів множини A перевищує кількість користувачів множини R), то користувач N_C отримує винагороду, а користувач N_R штрафується. Якщо $|A| < |R|$, то користувач N_R отримує винагороду за

неправомірну скаргу на нього, а N_C - штрафується. Відмітимо, що випадок $|A| = |R|$ неможливий з огляду на те, що кількість S суддів є непарним числом.

Передбачаємо, що усі користувачі множини U поділені на дві підмножини U_g та U_b ($U_g \cup U_b = U$).

$U_g \subset U$ - множина користувачів, які завжди поведуть себе коректно: користувач $u \in U_g$ створює лише коректний контент. Якщо такий користувач бере участь у судочинстві, скаргу на поганий контент він підтримає, скаргу на коректний контент він не підтримає.

$U_b \subset U$ - множина користувачів, які можуть поводити себе некоректно: користувач $u \in U_b$ може створювати як коректний, так і некоректний контент. Якщо такий користувач бере участь у судочинстві, скаргу на поганий контент він підтримає, скаргу на коректний контент він не підтримає. Якщо такий користувач бере участь у судочинстві, скаргу на поганий контент він не підтримає, скаргу на коректний контент він підтримає.

Нехай $|U_g| > |U_b|$, $|U_g|/|U_b| \geq k$, де k - коефіцієнт «допропорядності» учасників мережі.

Маємо вектор M електронних токенів («грошей») користувачів системи, який співвідноситься до множини користувачів U .

Користувачі без грошей не мають змоги створювати будь-який контент. Гіпотеза в тому, що така модель дозволить мінімізувати кількість грошей у користувачів, що створюють некоректний контент. .

2.3 Розробка алгоритму розв'язання

У системі інсуватиме два алгоритми - алгоритм скарги та алгоритм апеляції.

Формальний алгоритм скарги матиме таку послідовність:

КРОК 1. Створення скарги користувачем.

КРОК 2. Система обирає учасників I етапу консенсусу.

2.1 ЯКЩО більшість учасників приймає рішення щодо правомірності скарги ТО:

- 2.1.1 штрафувати того користувача, що створив контент;
- 2.1.2 надати нагороду користувачу, який створив скаргу;
- 2.1.3 надати нагороду всім суддям-учасникам рішення.

2.2 ЯКЩО більшість учасників приймає рішення щодо неправомірності скарги ТО:

- 2.2.1 штрафувати того користувача, що створив скаргу;
- 2.2.2 надати нагороду користувачу, який створив контент;
- 2.2.3 надати нагороду всім суддям-учасникам рішення.

Формальний алгоритм апеляції матиме таку послідовність:

КРОК 1. Створення апеляції користувачем, який був оштрафований на попередньому етапі.

КРОК 2. Система обирає учасників апеляційного етапу консенсусу серед користувачів, які не приймали участь в попередніх етапах.

2.1 ЯКЩО більшість учасників приймає рішення щодо правомірності апеляції ТО:

- 2.1.1 забрати винагороду у переможця попереднього етапу та оштрафувати його;
- 2.1.2 надати нагороду користувачу, що програв попередній етап;
- 2.1.3 надати нагороду користувачу, який ініціював апеляцію;
- 2.1.4 надати нагороду всім суддям-учасникам апеляційного.

2.2 ЯКЩО більшість учасників приймає рішення щодо неправомірності апеляції ТО:

- 2.2.1 оштрафувати особу, яка ініціювала апеляцію;
- 2.2.2 надати нагороду учасникам попереднього (правомірного) рішення;
- 2.2.3 надати нагороду всім суддям-учасникам апеляційного.

Маючи непарне число учасників у множині E , цей алгоритм завжди надаватиме консенсус у рішенні (тобто більшість або задовольнить скаргу, або не підтримає її). На цьому ж підході базуються і апеляції на прийняте рішення.

2.4 Опис імітаційної моделі

Задля перевірки моделі та алгоритму консенсусу було прийняте рішення про програмну реалізацію та створення імітаційної моделі за допомогою засобів мови C# та програмного фреймворку .Net Core. Було визначено наступні класи (об'єкти).

User - користувач системи, який приймає такі стратегії:

а) погані наміри - користувач створює допустимий і недопустимий контент; якщо він приймає участь у судочинстві - від підтримує поганий контент та голосує проти гарного;

б) гарні наміри - користувач створює тільки допустимий контент; якщо він бере участь у судочинстві - від підтримує гарний контент та голосує проти поганого. Кількість користувачів з гарними/поганими намірами визначається під час експеримента. Користувач також має прив'язаний до себе примітив класу Wallet.

Wallet - «гаманець» користувача з грошовими токенами. Якщо користувач створив контент, що визнано поганим, з цього рахунку списується фіксована сума. Розмір суми визначається під час експерименту. Якщо в користувача немає грошей – він не може створювати контент і приймати участь у судочинстві. Якщо користувач поскаржився на контент іншого користувача, та виграв скаргу, на його рахунок додається фіксована сума, яка дорівнює сумі, що списується у попередньому етапі.

Issue - скарга (або апеляція на скаргу), що пов'язується з користувачем, який створив контент, на який було створено скаргу, а також іншим користувачем, який цю скаргу створив. Скарга також вміщує в себе випадково обраних суддів, які й приймають рішення про її достовірність / недостовірність.

Якщо користувач вважає, що був оштрафований неправомірно, він може подати апеляцію. Подавши апеляцію, він витрачає додаткові гроші, що повині бути еквівалентні заставі, якщо цю апеляцію буде програно. Якщо апеляцію виграно, показники штрафів/винагород інвертуються, а також всі судді першого етапу, що проголосували за попереднє прийняте рішення, штрафуються.

Винагорода розподіляється між користувачем, що виграв апеляцію, та суддями наступного етапу

Модель має такі вхідні параметри:

- а) К - кількість користувачів системи;
- б) М - кількість токенів (грошей) у кожного користувача системи на початок моделювання;
- в) F - розмір винагороди/штрафу;
- г) J - кількість суддів, що обирається для розгляду скарги;
- д) Q - кількість повідомлень (на які можуть бути створено скарги), що поступають у систему у одиницю часу;
- е) Р - винагорода суддів за участь у розгляді справи; дорівнює 0 для спрощення

2.5 Результати експериментальних досліджень

Для експериментів визначено такі вхідні дані:

- а) кількість користувачів системи: 10 000 або 100 000;
- б) в початковий момент часу у кожного користувача 100 одиниць токенів (грошей);
- в) винагорода/штраф встановлені у розмірі 20 токенів;
- г) кількість суддів, що обираються для розгляду скарги, дорівнює 5;
- д) кількість повідомлень, що поступають в систему від користувачів, за одну секунду: 10 та 100 відповідно.

Для визначення значень параметрів системи консенсусу, які можуть забезпечити її ефективну роботу, було проведено ряд експериментів з різними вхідними параметрами залежно від відсотку «поганих» користувачів та наявності апеляцій в імітаційній моделі. Зокрема, були проведені експерименти з наступними вхідними параметрами:

- а) 5% поганих користувачів, без апеляцій;
- б) 5% поганих користувачів, з апеляціями;
- в) 30% поганих користувачів, без апеляцій;
- г) 30% поганих користувачів, з апеляціями;
- д) 45% поганих користувачів, без апеляцій.

Нехай P_b - сукупна кількість грошей у хороших користувачів, P_g - сукупна кількість грошей у поганих користувачів, а кількість суддів - 5. Для спрощення моделі, винагорода суддів встановлена такою, що дорівнює нулю. Метою моделювання було перевірка: чи виявить система некоректних користувачів і якщо так, то як швидко. Результати експерименту наведені в табл. 2.1. Вони свідчать, про те, що гіпотеза, яка перевіряється, вірна.

Таблиця 2.1 – Результати експериментального дослідження системи

Кількість користувачів (всього)	Грошей	Постів у секунду	$ U_g $	$ U_b $	Штраф/Винагорода	Ітерацій	Грошей (U_g)	Грошей (U_b)
10000	100	10	9500	500	20	2508	1000000	0
10000	100	10	7000	3000	20	6232	1000000	0
10000	100	10	9500	500	20	1793	1000000	0
10000	100	10	7000	3000	20	4976	1000000	0
10000	100	10	5500	4500	20	17350	1000000	0

В результаті проведених експериментів було виявлено наступне.

Чим більший відсоток поганих користувачів, тим більше ітерацій необхідно для остаточного перерозподілу коштів між хорошими і поганими користувачами. На рис. 2.1 відсоток поганих користувачів в системі становить 5% і на остаточний перерозподіл коштів необхідно 2508 ітерацій. На рис. 2.3 відсоток поганих користувачів в системі становить 30% і на остаточний перерозподіл коштів необхідно 6232 ітерацій. На рис. 2.5 відсоток поганих користувачів в системі становить 45% і на остаточний перерозподіл коштів необхідно 17350 ітерацій, однак, в даному випадку система веде себе непередбачувано.

Система є ефективною та передбачуваною при умові, що відсоток поганих користувачів не перевищує третину від загальної кількості користувачів системи. На рис. 2.1 та рис. 2.3 відсоток поганих користувачів менше третини і

система зберігає працездатність. На рис. 2.5 відсоток поганих користувачів перевищує третину і система веде себе непередбачувано.

Наявність у користувачів можливості подання апеляцій підвищує ефективність перерозподілу коштів між хорошими та поганими користувачами. Так, система що функціонує з допустимим відсотком поганих користувачів і підтримує процес подання апеляцій остаточно перерозподіляє кошти між хорошими і поганими користувачами за меншу кількість ітерацій порівняно з такою ж системою, що не підтримує процес подання апеляцій. Системі, що містить 5% поганих користувачів та підтримує апеляції (рис. 2.2) для остаточного перерозподілу коштів необхідно 1793 ітерації, тоді як такій же системі, що не підтримує апеляції (рис. 2.1), необхідно 2508 ітерацій. Системі, що містить 30% поганих користувачів та підтримує апеляції (рис. 2.4) для остаточного перерозподілу коштів необхідно 4976 ітерацій, тоді як такій же системі, що не підтримує апеляції (рис. 2.3), необхідно 6232 ітерації.

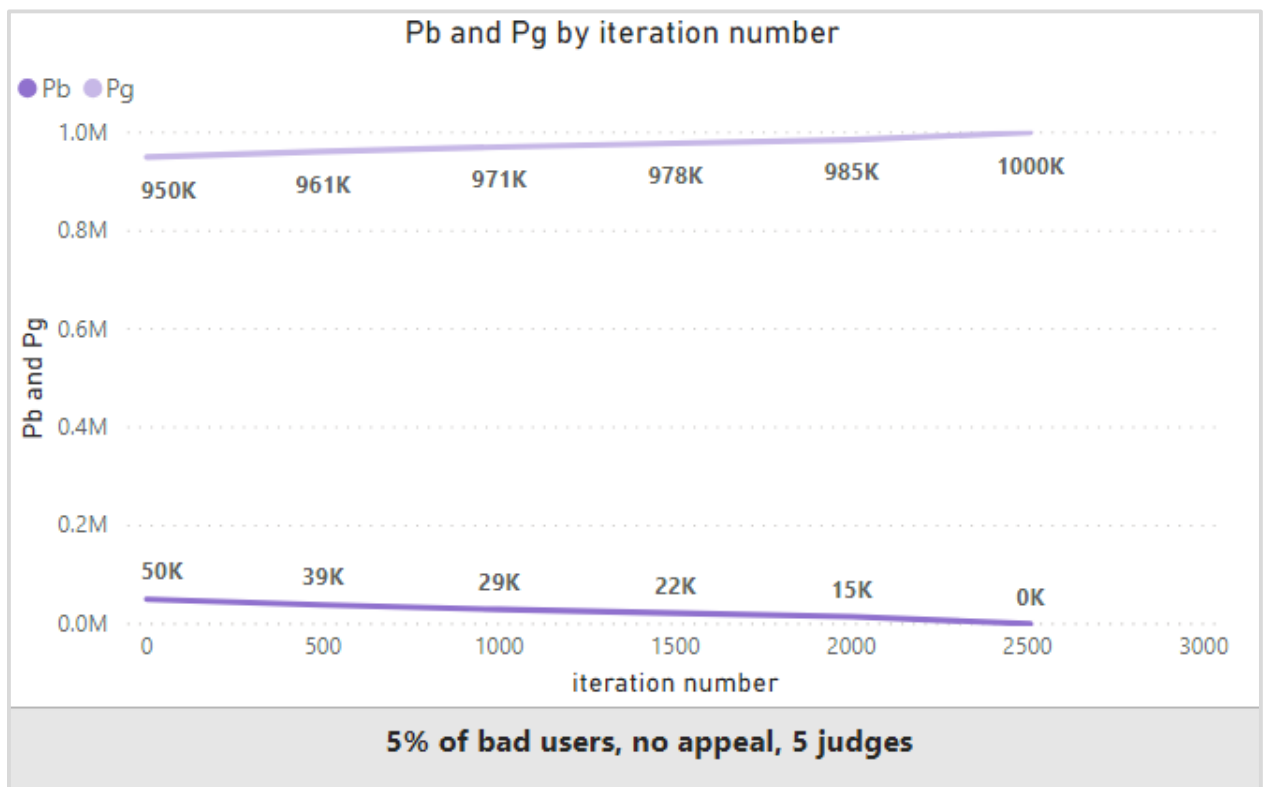


Рисунок 2.1 - Динаміка зміни об'єму грошей користувачів в часі при 5% поганих користувачів та відсутності апеляцій

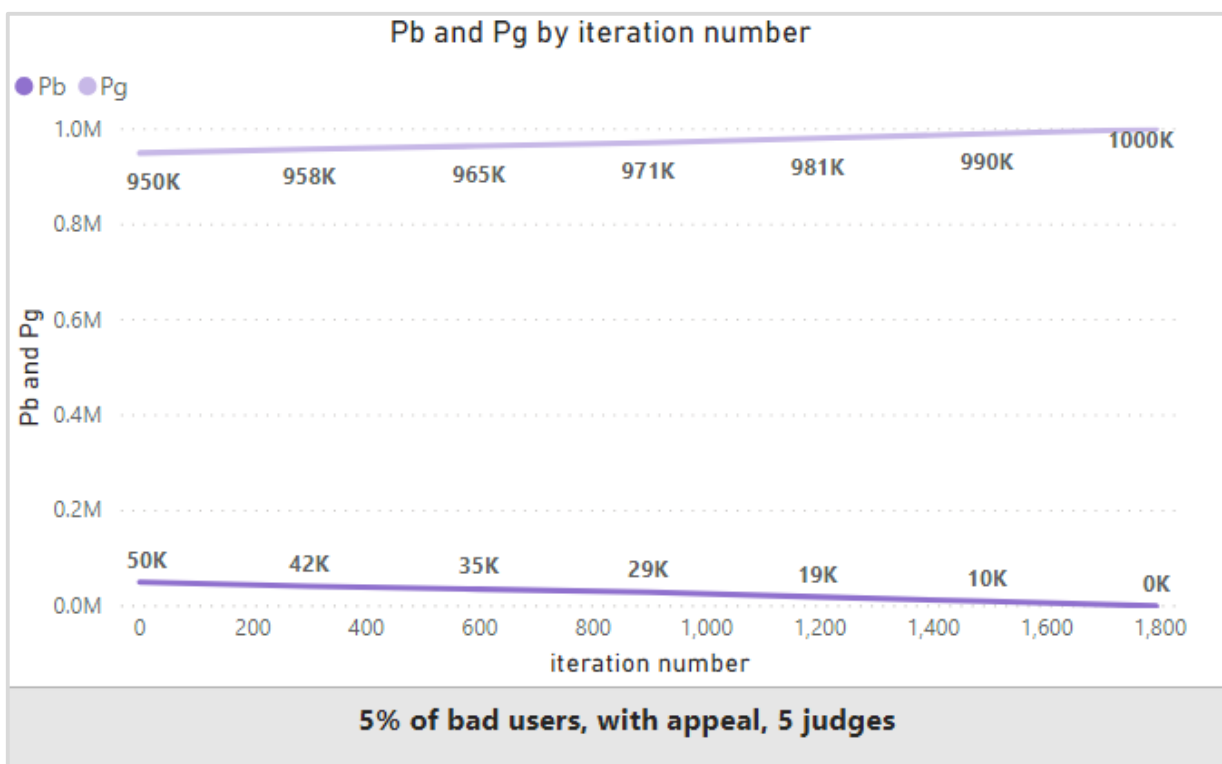


Рисунок 2.2 - Динаміка зміни об'єму грошей користувачів в часі при 5% поганих користувачів та наявності апеляцій

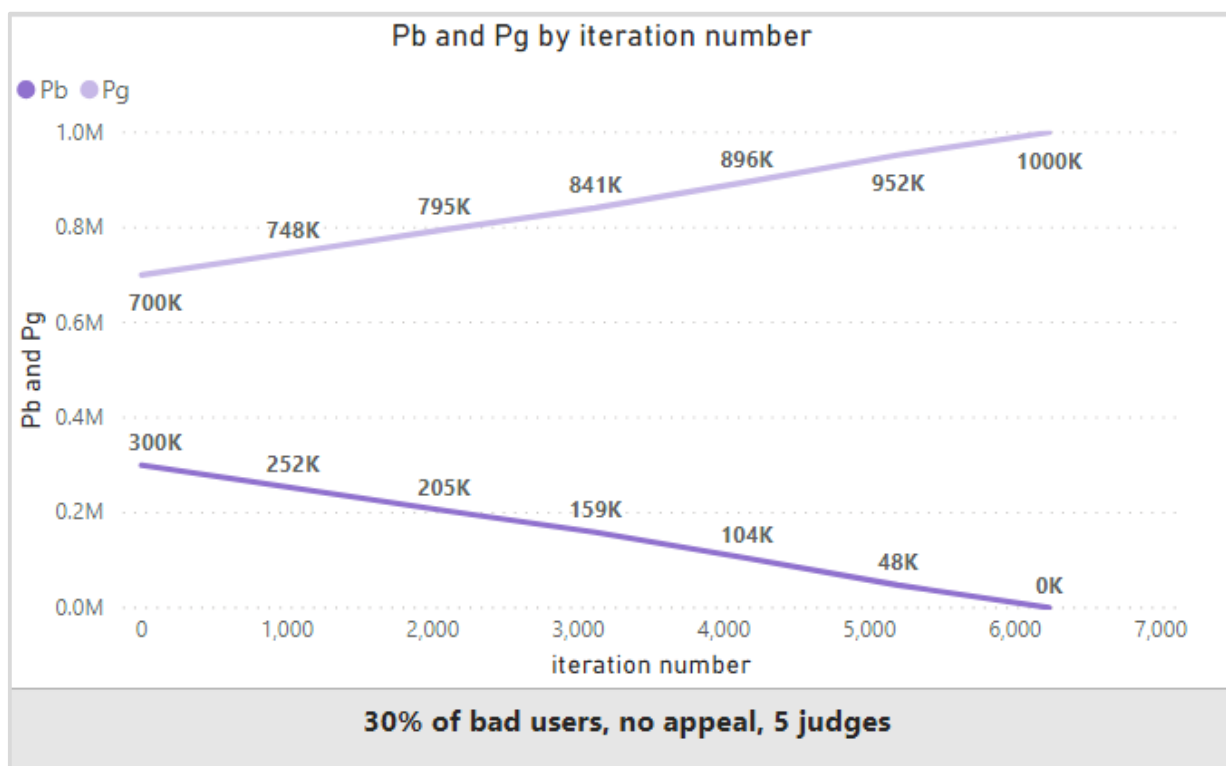


Рисунок 2.3 - Динаміка зміни об'єму грошей користувачів в часі при 30% поганих користувачів та відсутності апеляцій

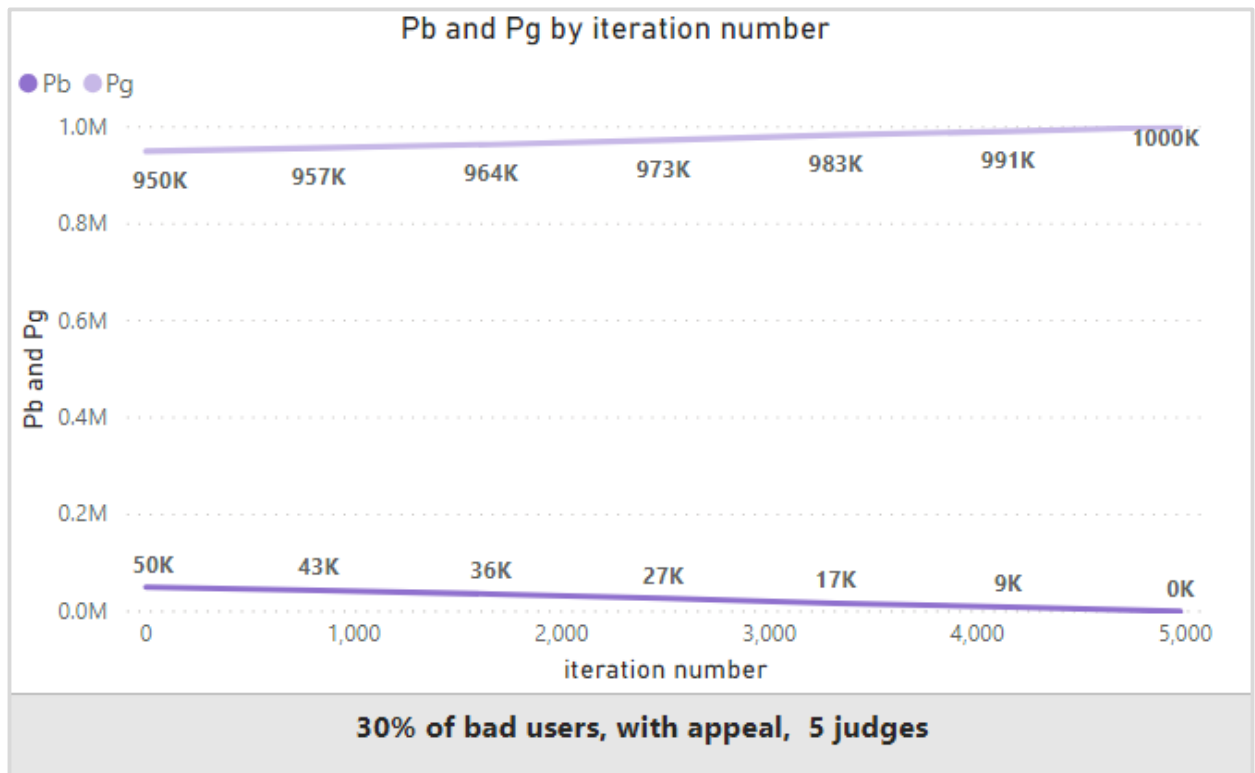


Рисунок 2.4 - Динаміка зміни об'єму грошей користувачів в часі при 30% поганих користувачів та наявності апеляцій

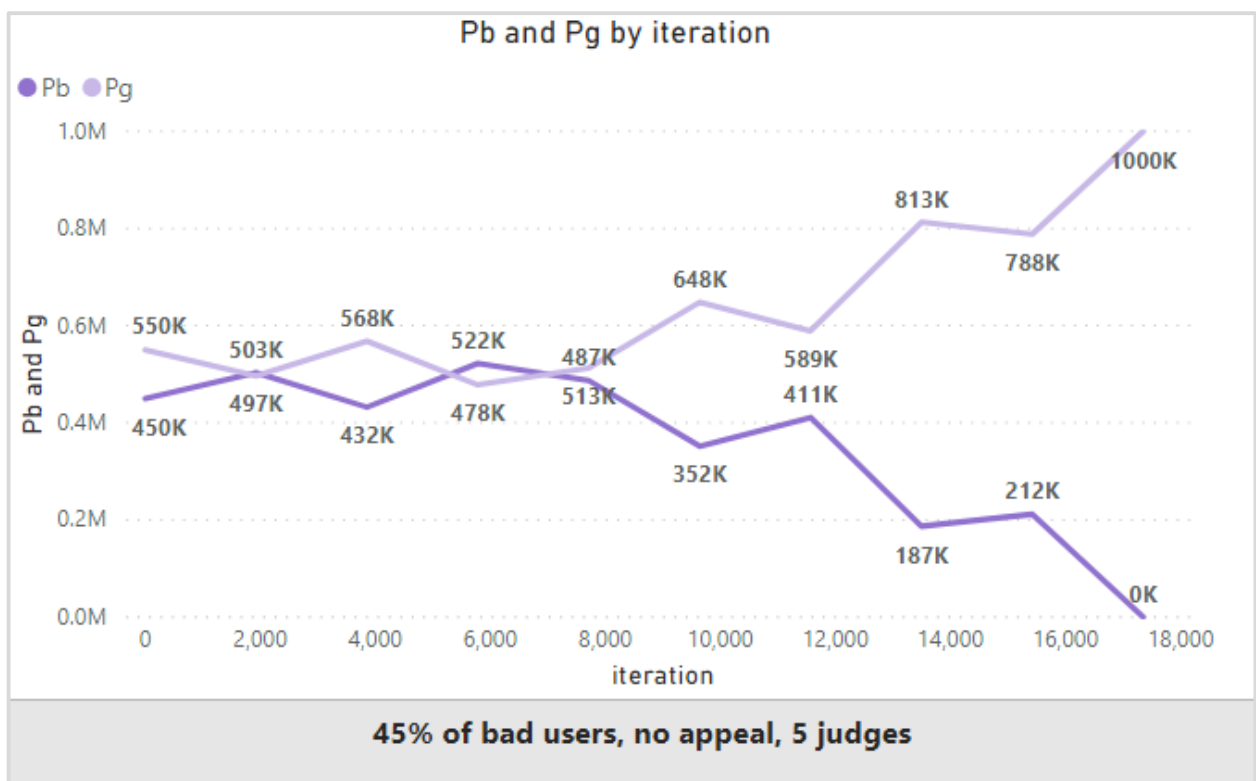


Рисунок 2.5 - Динаміка зміни об'єму грошей користувачів в часі при 45% поганих користувачів та відсутності апеляцій

Результати експериментів з різною кількістю суддів наведено в додатку А, плакат 4.

2.6 Висновок

В даному розділі описано вимоги до системи та концепцію алгоритму консенсусу. Також, створено математичну модель.

Алгоритм консенсусу буде мати скінчену кількість станів, а правила переходів користувачів у межах системи визначаються за допомогою умовних дерев рішень.

Розроблений алгоритм було перевірено експериментально, за допомогою створення імітаційної моделі: результати показали спроможність моделі консенсусу, за умови, що кількість «гарних» та кількість «поганих» користувачів суттєво відрізняється.

3 АНАЛІЗ ТЕХНОЛОГІЙ ДЛЯ ЗАБЕЗПЕЧЕННЯ ГОРИЗОНТАЛЬНОГО МАСШТАБУВАННЯ ДЕЦЕНТРАЛІЗОВАНОЇ МЕРЕЖІ

За результатами попереднього розділу, ми впевнились у спроможності запропонованого алгоритму консенсусу. Наступний виклик, з яким зтикаються інженери сучасних інформаційних систем - проектування та реалізація такої системи масового обслуговування, що може забезпечити потреби бізнесу. З огляду на те, що система орієнтована на спрощення проблем великої кількості користувачів (соціальні мережі, популярні блоги, можливо - громадське суспільство та законодавча влада), вважаю на необхідне інвестувати ресурси у створення правильної архітектури для цієї платформи. З викликами, що існують у реальних високонавантажених проектах, стикаються архітектори, технічні ліди та генеральні технічні директори по всьому світу. На даний час створено та пророблено велику кількість підходів, фреймворків та шаблонів, які вирішують різні класи проблем.

З огляду на те, що запропонована система дійсно не має аналогів у реальному світі, необхідно проаналізувати існуючі підходи у інших сферах діяльності (електронна комерція, великі дані, ін.). Виділено три основні технології, які можуть стати основою архітектури системи - патерн Saga для розподілених фінансових транзакцій, консистентне хешування для горизонтального масштабування та швидкісну нереляційну СКБД Apache Cassandra для зберігання даних, що часто змінюються.

3.1 Паттерн Saga

Операції є невід'ємною частиною програм. Без них неможливо було б підтримувати узгодженість даних. Одним з найпотужніших видів транзакцій є Two-Phase Commit (двофазне узгодження), що підсумовується, коли вчинення першої транзакції залежить від завершення другої. Це особливо корисно коли необхідно одночасно оновлювати кілька об'єктів, наприклад, як при підтвердженні замовлення. Однак при роботі з мікросервісами все ускладнюється. Кожна служба є окремою системою з власною базою даних, і

використання локальних двофазних узгоджень для підтримки послідовності всієї системи стає неможливим. Наприклад, інтернет-магазин, користувачам якого доступний кредитний ліміт, повинен гарантувати, що нове замовлення конкретного користувача не перевищує його кредитний ліміт. Так як «Замовлення» і «Клієнти» - різні бази даних, сервіс не може використовувати локальні ACID транзакції. В цьому випадку виникає проблема забезпечення узгодженості даних між сервісами. Для ілюстрації цієї проблеми розглянемо високорівневу мікросервісну архітектуру системи електронної комерції (рисунок 3.1).

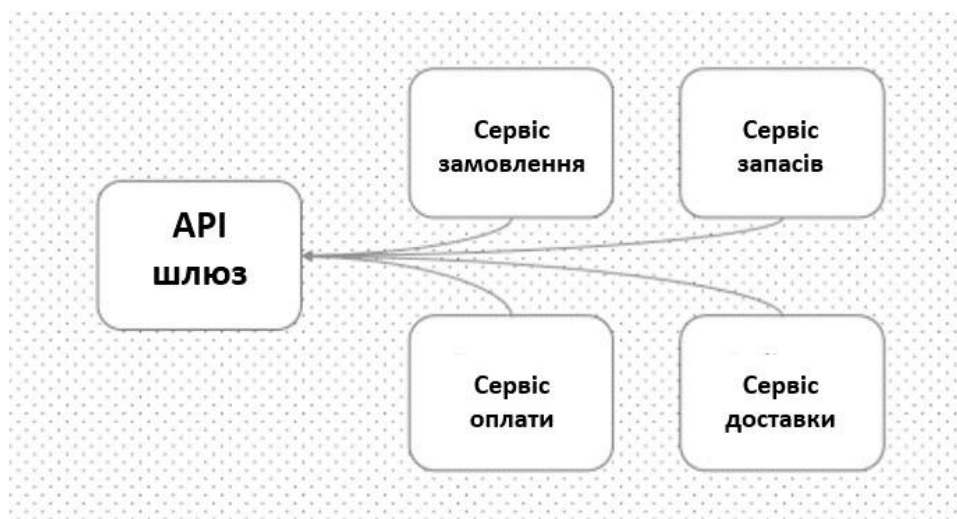


Рисунок 3.1 - Мікросервісна архітектура системи електронної комерції

У наведеному вище прикладі не можна просто оформити замовлення, стягнути кошти з клієнта, оновити запас та надіслати його для доставки в одній ACID-транзакції. Щоб послідовно виконувати весь цей потік, потрібно буде створити розподілену транзакцію. Одна з найбільш відомих моделей розподілених транзакцій називається Saga. Сага - це послідовність локальних транзакцій, в якому кожна транзакція оновлює дані в межах однієї послуги. Перша транзакція ініціюється зовнішнім запитом, що відповідає операції в системі, а потім кожен наступний крок ініціюється завершенням попереднього [16]. Кожна локальна транзакція оновлює базу даних і публікує повідомлення або подію ініціюючи наступну локальну транзакцію в межах саги. Якщо транзакція завершилася невдачею, наприклад, через порушення бізнес правил, то сага запускає компенсуючі транзакції, які «відкочують» зміни, зроблені

попередніми локальними транзакціями. Використовуючи попередній приклад електронної комерції високорівневий дизайн саги виглядатиме наступним чином (рисунок 3.2).

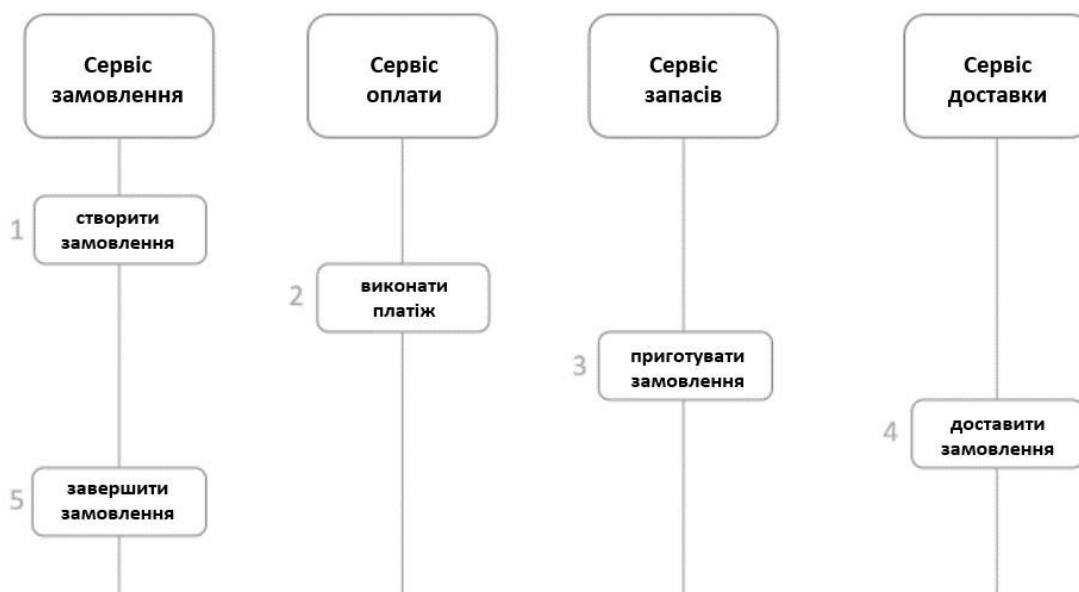


Рисунок 3.2 - Реалізація Saga в системі електронної комерції

Ідея, яку реалізує шаблон Saga, проста: після кожного успішно виконаного кроку ми маємо деякий стан, з якого можна буде продовжити виконання процесу. Кроком є виконання певної дії, реакція на подію тощо. Тобто, якщо ми не змогли підтвердити транзакцію в базі даних, або якщо виклик іншого веб-сервісу завершився невдачею, у нас є стан, що валідний на момент до його виклику. Бізнес-процес зупинений, але не втрачений і може бути відновлений за потреби.

Saga - це шаблон управління збоями. Модель Saga існує вже більше 30 років і широко використовується для управління розподіленими транзакціями. Saga використовується у розподілених системах по типу мікросервісних архітектур, які використовують одну базу даних на одну послугу (у випадку, коли анонімність не є необхідністю). По суті це один із способів збереження узгодженості даних для розподілених систем. При роботі з архітектурами, що базуються на мікросервісах, узгодженість даних часто стає проблемою для розробників. Saga з'явилася через усвідомлення того, що довготривалі транзакції і особливо далеко розподілені транзакції, що перетинають кордони розташування і/або довіри, не можуть бути оброблені з використанням класичної моделі ACID з двофазної фіксацією і утримуванням блокування на час роботи

(2PC). 2PC є протоколом двофазної фіксації транзакцій, який працює як одиночна комісія, спрямована на здійснення транзакцій ACID в розподілених системах. Він використовується там, де важлива стійка послідовність. Saga, натомість, працює послідовно, а не як окрема комісія. Кожна операція виконується перед наступною операцією, що в кінцевому підсумку робить дані узгодженими. Saga складається з декількох етапів, тоді як 2PC діє як один запит (рисунок 3.3).

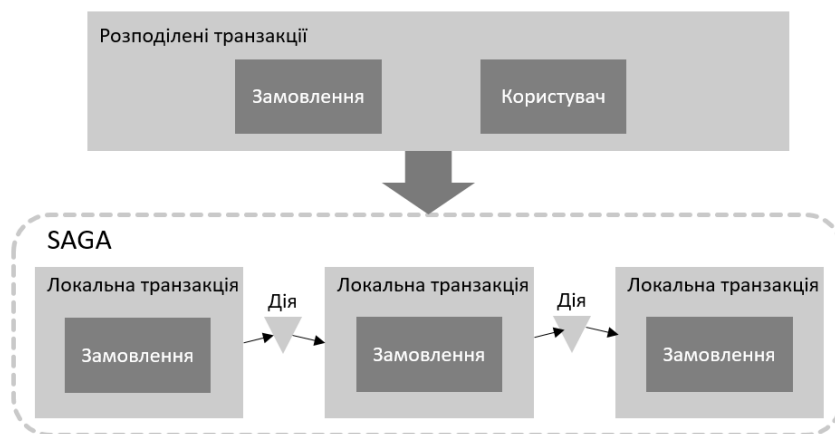


Рисунок 3.3 - Використання Saga замість 2PC

Таким чином, Saga розбиває роботу на окремі транзакції, результати яких можуть бути певним чином скасовані після того, як робота виконана. Наприклад, при бронюванні відпустки онлайн, система спочатку повертає повідомлення про успішне бронювання користувачу, а потім послідовно виконує бронювання рейсу, бронювання готелів, бронювання автомобілів тощо, використовуючи різні підсистеми. На рисунку 3.4 зображений шаблон Saga в дії.

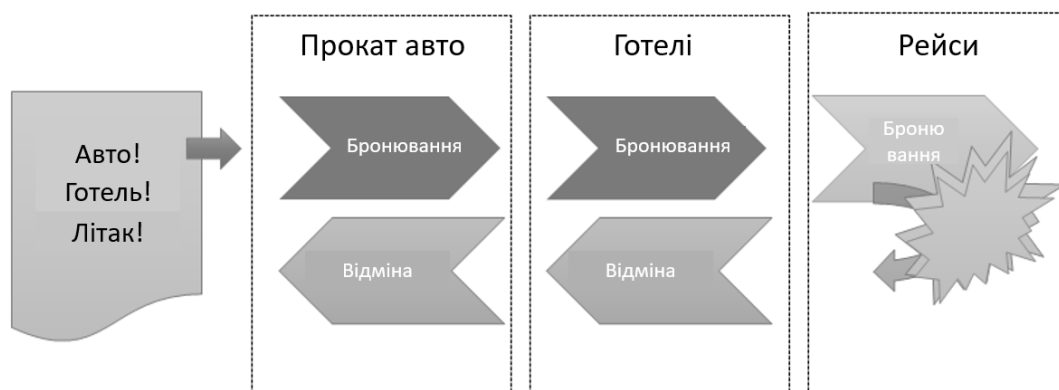


Рисунок 3.4 - Приклад застосування шаблону Saga

Якщо ви бронюєте туристичний маршрут, вам потрібен автомобіль, готель і рейс. Якщо ви не можете отримати все з переліченого, поїздка, ймовірно, не

зможеться відбутися. Також важливим є те, що ви не можете залучити всіх постачальників перелічених послуг до розподіленої транзакції ACID. Замість цього у вас будуть активності з бронювання автомобіля, готельного номеру і квитка на літак, які знають, як виконувати і скасовувати бронювання. Дії групуються в складене завдання (маршрутний лист), а це завдання передається по ланцюжку дій. Якщо дія завершується успішно, вона додає запис про завершення до маршрутного листа разом з інформацією про те, де може бути досягнута компенсаційна операція. Якщо дія завершується невдало, вона очищається локально, а потім відправляє прокрутку/відкат маршруту назад, крізь всі елементи списку до останньої адреси компенсації, щоб скасувати результат транзакції [17].

Існує кілька способів координації саг, але найбільш популярними є два:

- а) Choreography (Хореографія);
- б) Orchestration (Оркестровка).

Хореографія. При цьому способі кожна транзакція публікує події, які запускають транзакції в інших сервісах [18]. Центральної координації немає, кожна служба виробляє та прослуховує події інших служб і вирішує, чи слід вжити заходів. При такому підході в системі не було б оркестратора. Кожна служба виконує свій власний процес, і якщо результат є успішним, він надсилає відгук про успіх для наступного кроку. У разі відмови він надсилає відгук про відмову для попереднього кроку. Таким чином, служби, які працювали до цього кроку, можуть послідовно виконати відкати. Хореографія - це природний спосіб втілити схему саги, вона проста, зрозуміла, не потребує великих зусиль для побудови, а всі учасники нещільно пов'язані, оскільки вони не мають прямого знання один про одного. Якщо транзакція передбачає 2 - 4 кроки, це може бути вдалим рішенням. У випадку більшої кількості кроків події та дизайн можуть бути складнішими для управління. Такий підхід може швидко стати заплутаним при додаванні додаткових кроків в транзакцію, оскільки важко відстежити, які служби які події прослуховують. На рисунку 3.5 зображена координація саг на основі хореографії з використанням подій.

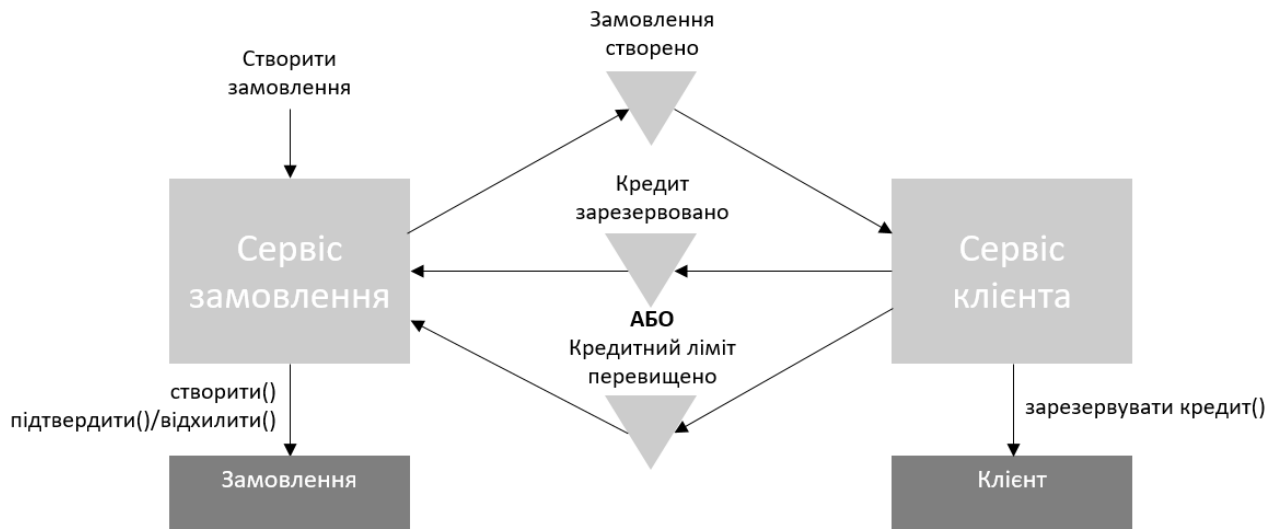


Рисунок 3.5 - Saga, заснована на хореографії

В інтернет магазині з використанням саги, заснованої на хореографії, створення замовлення буде включати наступні кроки:

- а) сервіс замовлення створює замовлення зі статусом «в очікуванні» і публікує подію «замовлення створено»;
- б) сервіс клієнта отримує подію і намагається зарезервувати кредит для замовлення, після чого публікує одну з двох подій: «кредит зарезервовано» або «кредитний ліміт перевищено»;
- в) сервіс замовлення отримує подію і переводить замовлення в стан «підтверджено» або «скасовано».

Оркестровка. При цьому способі оркестратор говорить учасникам, які транзакції повинні бути запуснені [18]. Оркестратор відповідає за централізацію прийняття рішень саги та послідовність логіки бізнесу. На рисунку 3.6 зображена координація саг на основі оркестровки.

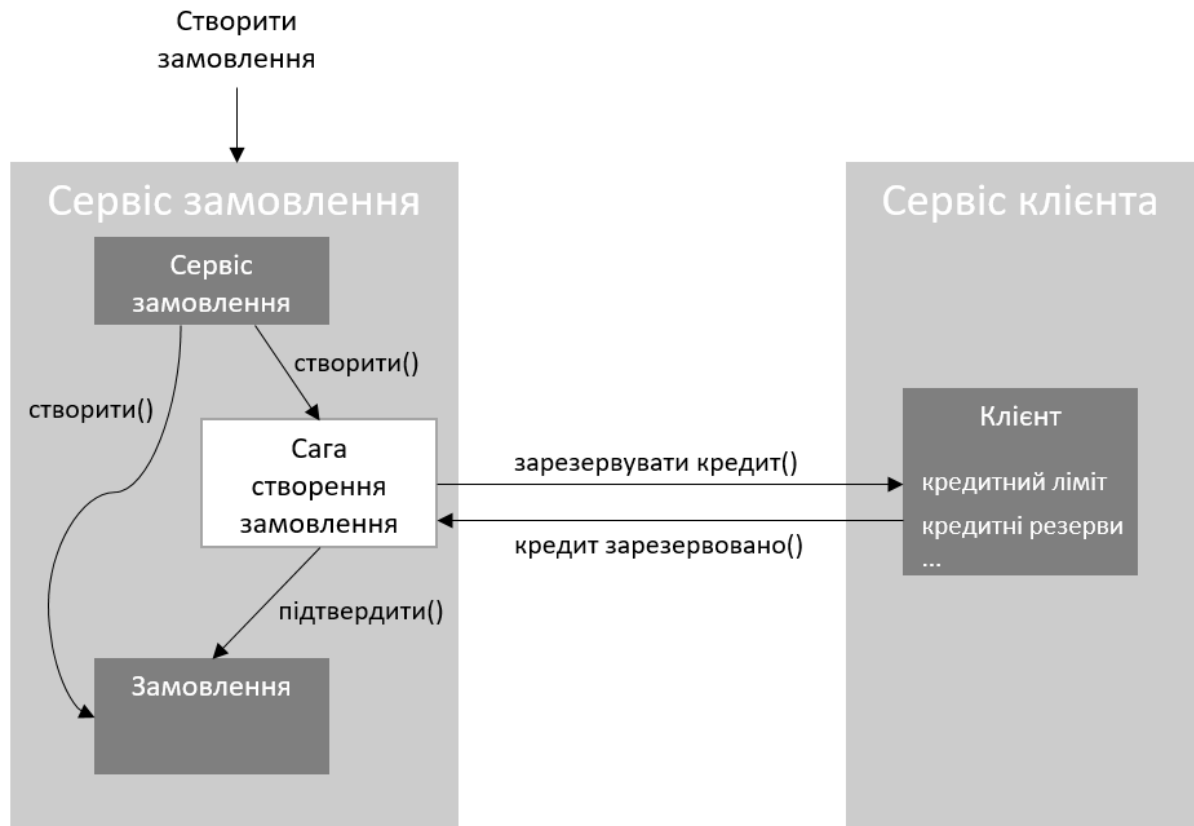


Рисунок 3.6 - Сага, заснована на оркестровці

В інтернет магазині з використанням саги, заснованої на оркестровці, створення замовлення буде включати наступні кроки:

- а) сервіс замовлення створює замовлення зі статусом «в очікуванні» і створює сагу створення замовлення;
- б) сага створення замовлення відправляє команду «зарезервувати кредит» в сервіс клієнта;
- в) сервіс клієнта намагається зарезервувати кредит для замовлення і відправляє назад відповідь;
- г) сага створення замовлення отримує відповідь і відправляє команду «підтвердити замовлення» або «відмінити замовлення» в сервіс замовлення;
- д) сервіс замовлення переводить замовлення в стан «approved «підтверджено» або «скасовано».

Основна перевага шаблону Saga полягає в тому, що він допомагає підтримувати узгодженість даних між сервісами без використання розподілених транзакцій. Це надзвичайно важливий аспект для архітектури мікросервісів. Однак головним недоліком шаблону Saga є складність з точки зору програмування. Також розробники не настільки звикли до написання саг

порівняно з написанням традиційних транзакцій. Інша проблема полягає в тому, що компенсаційні транзакції також повинні бути розроблені для того, щоб працювати з сагами [19].

3.2 Консистентне хешування

Перед тим, як розглядати консистентне хешування, варто ознайомитись з базовими поняттями, що стосуються хешування даних. Хеш-таблиця за своєю суттю є масивом, що працює з хеш-функцією. Хеш-функція - це функція, яка з'єднує один фрагмент даних (зазвичай той, що описує якийсь об'єкт, часто довільного розміру) з іншим фрагментом даних (зазвичай типовим цілим числом, що відоме як хеш-код, або просто хеш). Хеш-функція отримує вхідну частину даних, що називається ключем, а на виході видає ціле число, що називається хеш-значенням (рисунок 3.7).

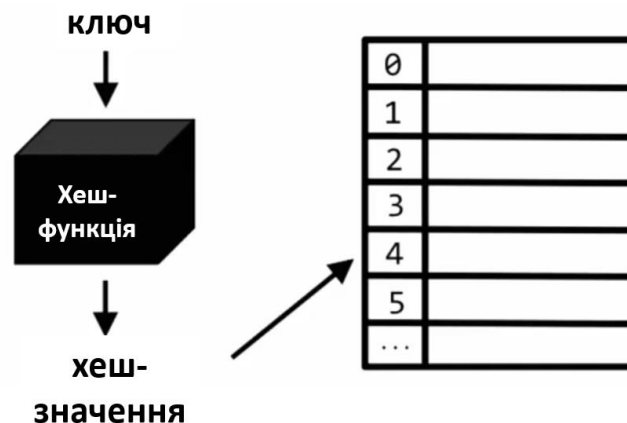


Рисунок 3.7 - Алгоритм роботи хеш-функції

Хеш-значення прив'язує ключ до певного індексу хеш-таблиці. Спочатку ми використовуємо хеш-функцію, щоб визначити, де саме в таблиці буде збережений певний ключ (записати). Потім ми використовуємо ту ж саму функцію, щоб знайти в таблиці збережений нами ключ (знайти). Хеш-функція може використовуватись для збереження будь-яких типів даних. Наприклад, деяка хеш-функція, призначена для хешування символічних рядків з діапазоном вихідних значень від 0 до 24, може з'єднувати рядок «apple», скажімо, з числом 0, а «banana» - з числом 1. В наведеному прикладі хеш функція приймає на вхід слово і видає на вихід певним чином сконвертоване у числове значення вхідне слово (рисунок 3.8).

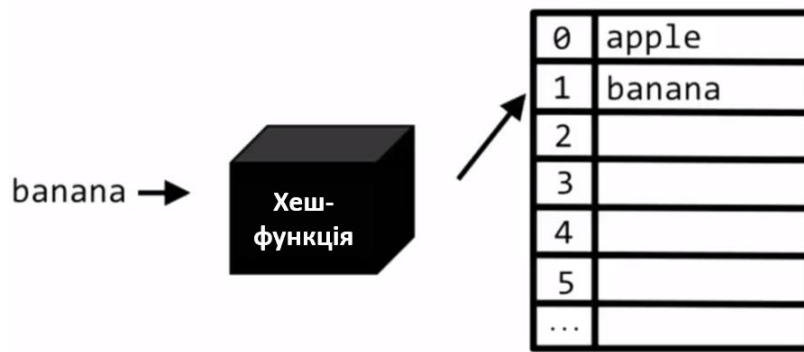


Рисунок 3.8 - Приклад роботи хеш-функції

Дана хеш-функція конвертує символічні значення у числові відповідно до їх першого символу (його порядкового номеру в алфавіті). Оскільки множина можливих входів більша за множину можливих виходів, будь-яке задане число матиме багато різних рядків, що йому відповідають. Це явище носить назву колізія. Наприклад, слово «ant» хешується функцією так же, як і слово «apple» - два ключі з'єднуються з одним індексом (рисунок 3.9).

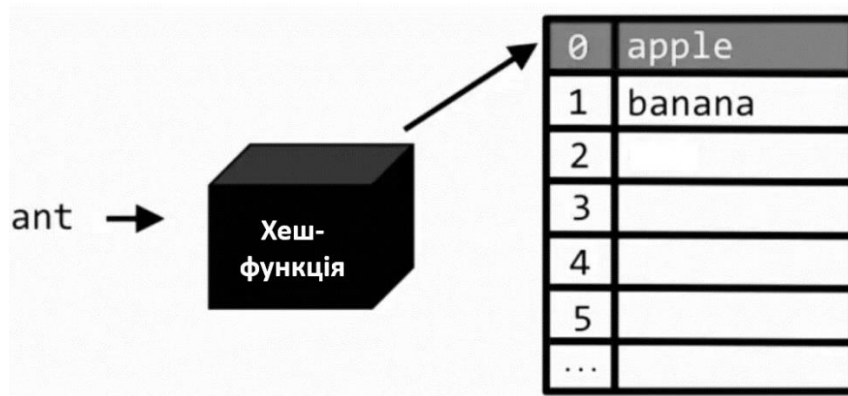


Рисунок 3.9 - Приклад колізії в хеш-таблиці

Найбільш застосовуваним методом вирішення колізій є метод ланцюжків (рисунок 3.10).

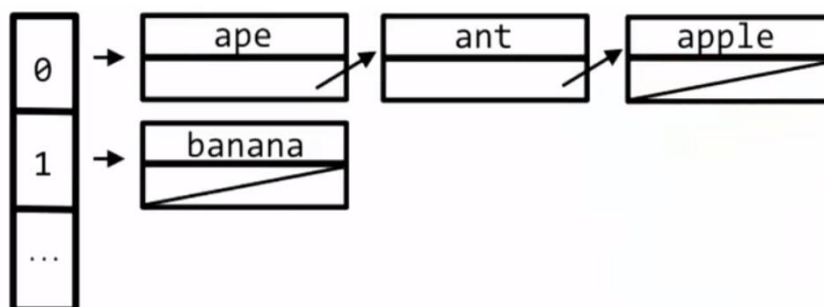


Рисунок 3.10 - Метод ланцюжків для вирішення колізій

В цьому випадку хеш-таблиця стає масивом вказівників на зв'язні списки. Тобто, якщо при додаванні нового значення виникає колізія, ключ додається в

початок зв'язного списку відповідного індексу. Таким чином, в разі пошуку в таблиці слова «apple» в найгіршому випадку необхідно буде пройти весь зв'язний список починаючи з індексу 0. Вирішення колізій приводить до того чи іншого типу лінійного пошуку, який значно сповільнює пошук. Тому в першу чергу потрібно використовувати таку функцію, яка мінімізує шанс виникнення колізій. Хороші хеш-функції повинні хешувати вхідні дані таким чином, щоб виходи для різних вхідних значень розподілялися максимально рівномірно у межах виходу. Це в свою чергу дозволить зробити довжину зв'язних списків приблизно однаковою у випадку виникнення колізій.

Класичний підхід хешування використовує хеш-функцію для генерування псевдовипадкового числа, яке потім ділиться на розмір простору пам'яті щоб перетворити випадковий ідентифікатор в положення у доступному просторі пам'яті, як зображено на рисунку 3.11.

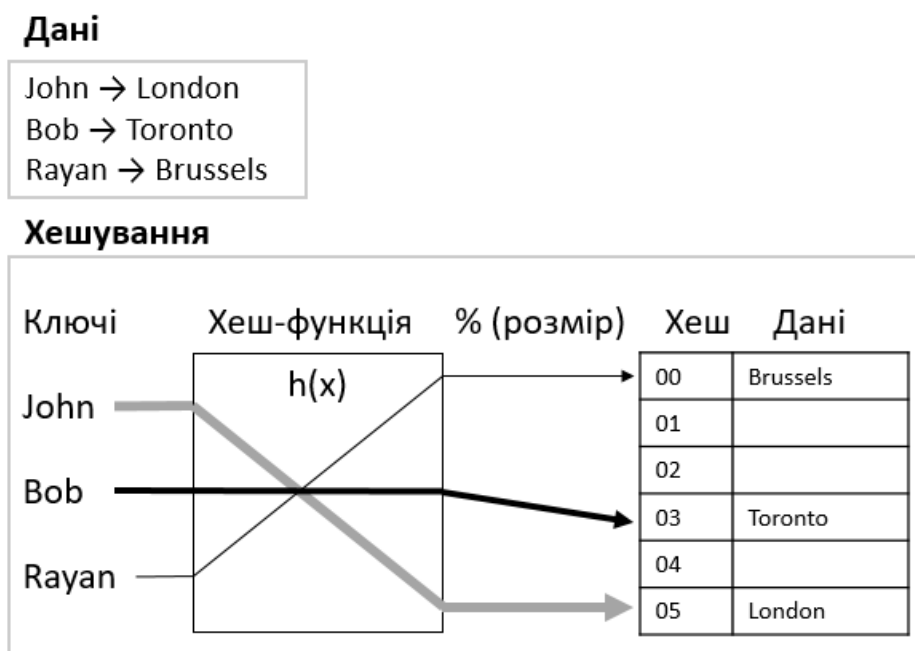


Рисунок 3.11 - Приклад класичного хешування

У випадку, коли різні програми, комп'ютери або користувачі потребують певних ресурсів з декількох вузлів сервера, нам потрібен механізм щоб рівномірно розподіляти запити на доступні серверні вузли, таким чином забезпечуючи збалансованість завантаження та підтримку стабільної продуктивності. Ми можемо розглядати вузли сервера як заповнювачі, до яких

можуть бути під'єднані один або більше запитів. В класичному методі хешування ми завжди припускаємо, що:

- а) відома кількість місць пам'яті;
- б) кількість місць пам'яті ніколи не змінюється.

Сервери можуть масштабувати розмір кластера вгору і вниз протягом дня паралельно ліквідуючи несподівані збої. Але, якщо врахувати вищезазначений сценарій, неможливо гарантувати, що кількість вузлів сервера залишиться однаковою. Що робити, якщо один з вузлів перестане працювати? При класичному хешуванні у випадку відмови одного з вузлів потрібно повторно перехешувати кожен ключ, оскільки нове з'єднання «ключ - хеш-значення» залежить від кількості вузлів/місць пам'яті, як зображено на рисунку 3.12.

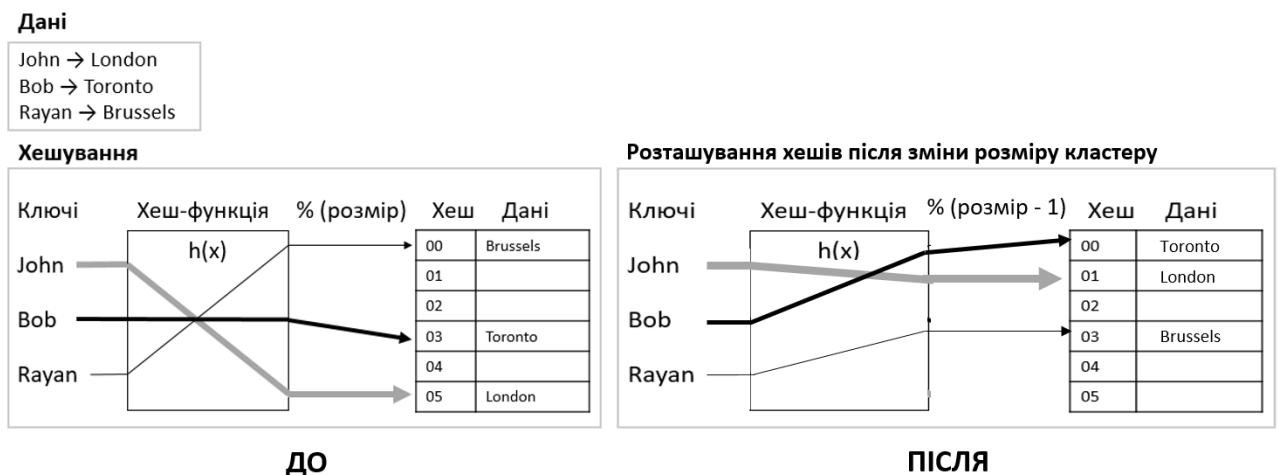


Рисунок 3.12 - Зміна розташування хешів при класичному хешуванні

Проблема класичного хешування в розподіленій системі - куди переміщається місце розташування кожного ключа (куди перехешується кожен ключ). Вона полягає в тому, що у кожного вузла є стан, який зберігається. Наприклад, незначна зміна розміру кластера може спровокувати величезну кількість роботи для здійснення перестановки всіх даних навколо кластера. Зі збільшенням розміру кластера цей процес стає нестійким, оскільки кількість роботи, необхідної для кожної зміни хешу, лінійно зростає з розміром кластера. Саме тут виникає необхідність у консистентному хешуванні.

Консистентне хешування є способом створення розподілених хеш-таблиць, при якому падіння одного або більше серверів-сховищ не призводить до необхідності повного перерозміщення всіх х1ключів і значень. Консистентне

хешування - це схема розподіленого хешування, яка працює незалежно від кількості серверів або об'єктів у розподіленій хеш-таблиці шляхом призначення їм позицій на абстрактному колі (хеш-кільці) [20]. Це дозволяє серверам та об'єктам масштабуватись, не впливаючи на всю систему. Консистентне хешування полегшує розподіл даних по набору вузлів сервера таким чином, щоб мінімізувати повторне з'єднання/ перевизначення даних при додаванні нових або видаленні існуючих вузлів. Консистентне хешування є ефективним прийомом, що використовується для внесення масштабованості в архітектуру сховища системи з самого початку. У розподіленій системі консистентне хешування допомагає вирішити наступні сценарії:

а) забезпечити еластичність масштабування (термін, що використовується для опису динамічного додавання/видалення серверів на основі завантаженості) для серверів;

б) забезпечити масштабування набору вузлів зберігання даних, таких як NoSQL бази даних.

Консистентне хешування можна описати наступним чином. Воно представляє запитувачів ресурсів і серверні вузли у вигляді певної віртуальної кільцевої структури, що носить назву «hashring» (з англ. hash - хеш, ring - кільце) - хеш-кільце. Кількість місць не фіксована, але вважається, що кільце має нескінченну кількість точок, і вузли сервера можуть розміщуватися у випадкових місцях на цьому кільці. Звичайно, можна знову вибрати це випадкове число за допомогою хеш-функції, але другий крок ділення його на кількість доступних місць пропускається, оскільки це вже не є скінченним числом. Запитувачі ресурсів, тобто, користувачі, комп'ютери або безсерверні програми, які є аналогами ключів у класичному хешуванні, також розміщуються на одному кільці, використовуючи ту саму хеш-функцію (рисунки 3.13).

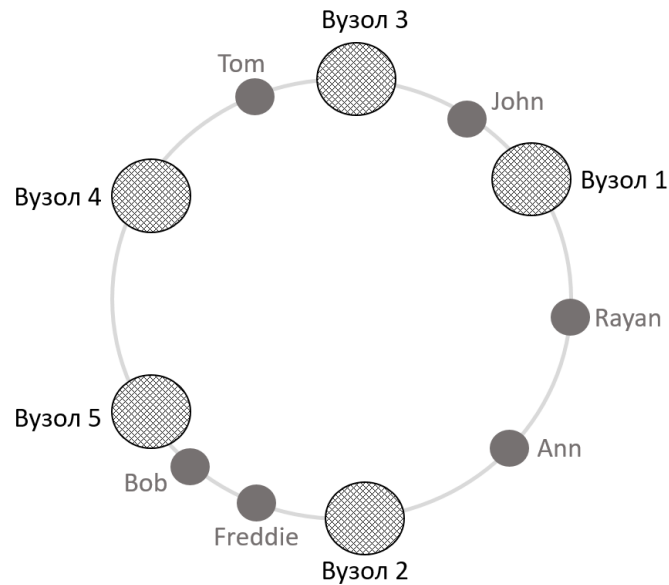


Рисунок 3.13 - Хеш-кільце

Отже, як приймається рішення про те, який запит яким серверним вузлом буде обслуговуватися? Якщо припустити, що кільце впорядковано таким чином, щоб обхід кільця за годинниковою стрілкою відповідає порядку зростання адрес локацій, кожен запит може обслуговуватися вузлом, який вперше з'являється в цьому обході за годинниковою стрілкою. Тобто, перший серверний вузол з адресою, більшою за адресу запиту, отримує його для обслуговування. Якщо адреса запиту вище, ніж найвищий адресований вузол, він обслуговується вузлом сервера з найменшою адресою, оскільки проходження через кільце здійснюється по колу [20]. Це проілюстровано на рисунку 3.14.

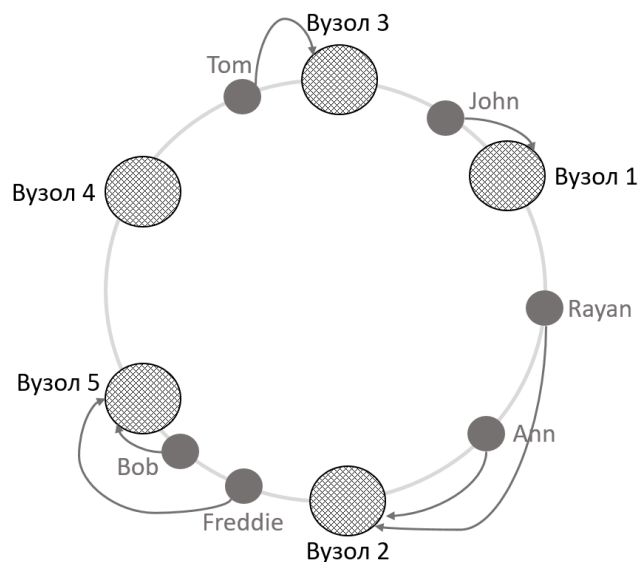


Рисунок 3.14 - З'єднання в хеш-кільці

Теоретично, кожен вузол сервера «володіє» діапазоном кільця хешів, і будь-які запити, що надходять в межах цього діапазону, будуть обслуговуватися

відповідним вузлом сервера. Якщо один з вузлів сервера виходить з ладу, скажімо, Вузол 3, діапазон наступного вузла розширюється і будь-який запит, що надходить в межах всього цього діапазону, переходить до новоствореного серверного вузла. В цьому випадку є лише один діапазон хешів, які підлягають повторному перепризначенню, тоді як решта діапазонів кільця хешів і призначень «вузол – запит» залишаються незмінними. Це головна відмінність від класичної моделі хешування, в якій зміна розміру хеш-таблиці порушує всі з'єднання. Завдяки консистентному хешуванню певна зміна кільця вплине лише на частину (відносно до коефіцієнта розподілу кільця) запитів. Зміна кільця відбувається через додавання або видалення вузла, що призводить до зміни деяких з'єднань «вузол - запит» (рисунок 3.15).

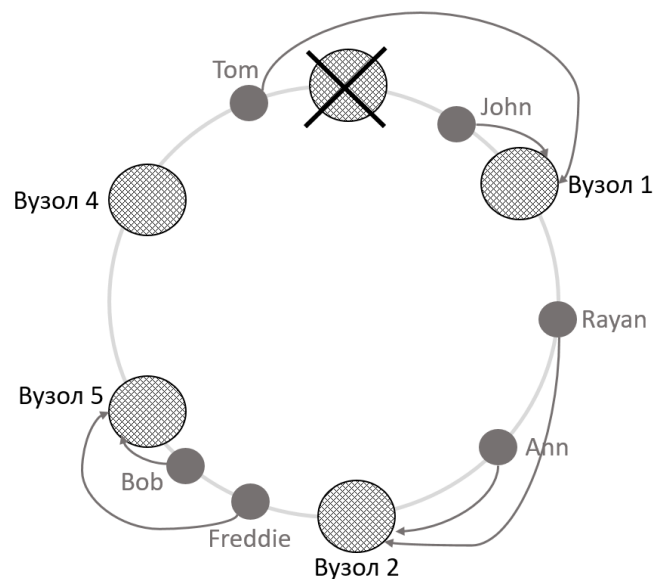


Рисунок 3.15 - Вихід з ладу вузла №3 в кільці хешів

Розглянемо алгоритм консистентного хешування більш детально на конкретному прикладі [21]. Нехай необхідно створити простір хеш-ключів. Нехай h є функція, яка генерує цілі числові хеш-значення в діапазоні $[0, 2^{32}-1)$. Можна представити це як масив цілих чисел з $2^{32} - 1$ слотами. Перший слот - h_0 , останній слот - h_{n-1} (рисунок 3.16).



Рисунок 3.16 - Простір хеш-ключів

Нехай хеш-простір представляється у вигляді кільця. Тобто, попередньо згенеровані цілі числа розміщуються на кільці таким чином, що останнє значення огортає кільце (рисунок 3.17).

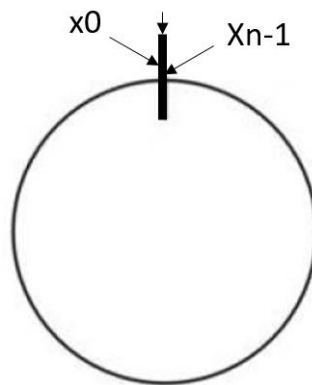


Рисунок 3.17 - Візуалізація хеш-простору в якості кільця

Нехай в даному кільці мають бути розміщені сервери баз даних і є список серверів, з яких слід почати. Використовуючи хеш-функцію, ми з'єднуємо кожен сервер бази даних з певним місцем на хеш-кільці. Наприклад, якщо у нас є 4 сервери, ми можемо використовувати хеш їх IP-адрес, щоб з'єднати їх з різними цілими числами за допомогою хеш-функції. Це імітує розміщення чотирьох серверів в різних місцях кільця (рисунок 3.18).

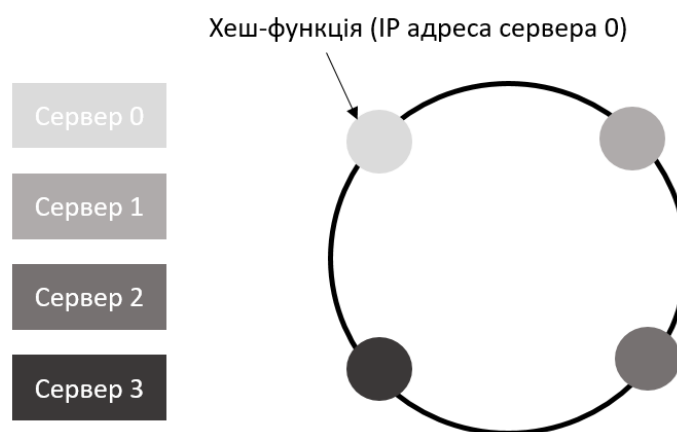


Рисунок 3.18 - Розташування серверів баз даних на хеш-кільці

Нехай потрібно визначити розташування ключів на даних серверах. Щоб знайти, на якому сервері знаходиться вхідний ключ (щоб додати його чи зробити запит на нього), треба зробити наступне:

а) пропустити ключ через ту саму хеш-функцію, яка використовувалась для визначення розташування сервера на кільці;

б) після хешування ключа отримати ціле значення, яке буде міститися в хеш-просторі, тобто, воно може бути з'єднано з деякою позицією в хеш-кільці. Тут може бути два випадки.

Перший випадок. Хеш-значення з'єднується з місцем кільця, у якого немає відповідного сервера. У цьому випадку ми рухаємося за годинниковою стрілкою на кільці від точки, з якою з'єднаний ключ, до першого сервера, що зустрінеється на кільці. Як тільки ми знайдемо перший сервер, подорожуючи по кільцю за годинниковою стрілкою, вставляємо туди ключ. Та сама логіка буде застосовуватись при спробі знайти ключ на кільці.

Другий випадок. Хеш-значення ключа з'єднується безпосередньо з тим же хеш-значенням сервера - в такому випадку ключ розміщується на даному сервері.

Наприклад, нехай є 4 вхідні ключі: ключ 0, ключ 1, ключ 2, ключ 3, і жоден з них безпосередньо не відображає хеш-значення будь-якого з 4 серверів хеш-кільця. Таким чином, для кожного ключа необхідно рухатись по кільцю за годинниковою стрілкою від точки, з якою з'єднується ключ, до першого серверу, куди і вставляється ключ (рисунок 3.19).

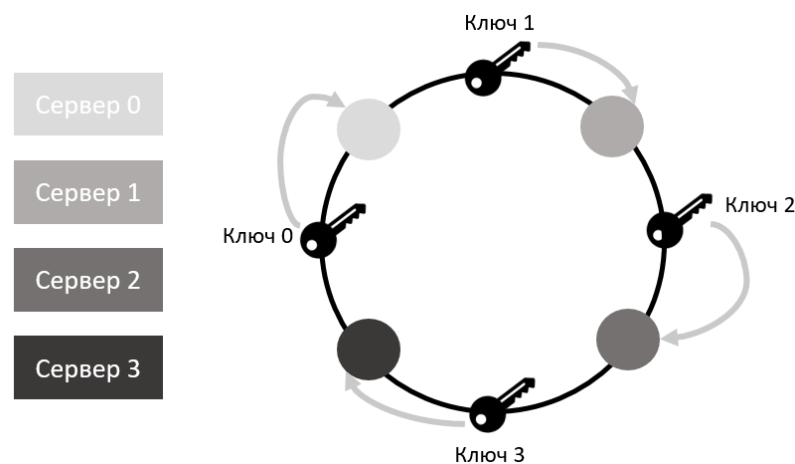


Рисунок 3.19 - Розміщення ключів на серверах в хеш-кільці

Нехай необхідно додати сервер до кільця. При додаванні до хеш-кільця нового сервера (сервер 4) потрібно буде перез'єднати ключі. Однак, перез'єднати з сервером 4 необхідно лише ті ключі, що знаходяться в діапазоні між сервером 3 і сервером 0. В середньому, потрібно буде перез'єднати k/n ключів, де k - кількість ключів, а n - кількість серверів. Локальне перез'єднання ключів є головною перевагою порівняно з глобальним перез'єднанням всіх ключів у

класичному хешуванні. На рис. зображений ефект від додавання нового сервера 4 - оскільки сервер 4 знаходиться між ключем 0 та сервером 0, ключ 0 буде перепід'єднаний від сервера 0 до сервера 4 (тобто, до першого сервера, який зустрічається на оновленому кільці якщо рухатись за годинниковою стрілкою від поточного ключа) (рисунок 3.20).

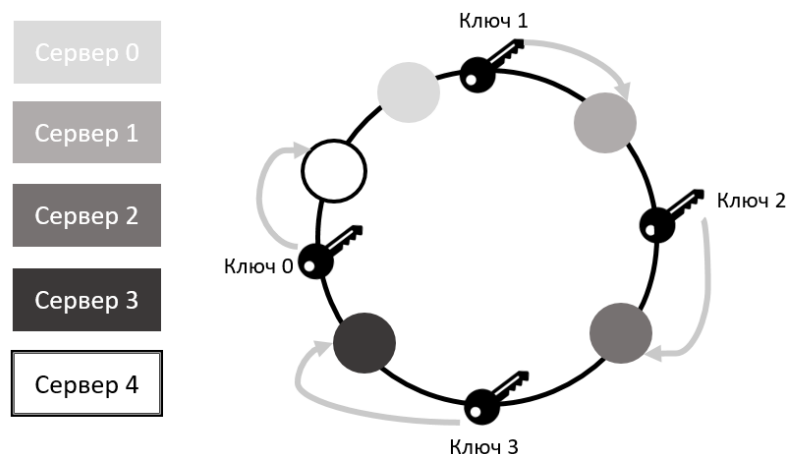


Рисунок 3.20 - Ефект додавання серверу до хеш-кільця

Нехай необхідно видалити існуючий сервер з кільця. Кількість серверів може зменшитись в процесі роботи, і послідовна схема хешування забезпечує мінімальний вплив на кількість ключів та серверів, які потребують змін. Якщо сервер 0 зазнає збоїв, лише ключі, що знаходяться в діапазоні між сервером 3 та сервером 0, будуть перепід'єднані до серверу 1, а решта ключів залишаться без змін (рисунок 3.21).

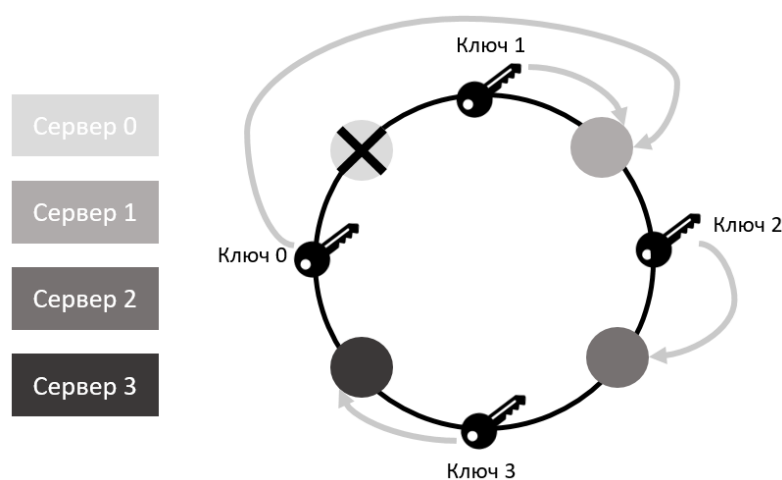


Рисунок 3.21 - Ефект видалення серверу з хеш-кільця

На цьому етапі зрозуміло, що консистентне хешування успішно вирішує проблему горизонтальної масштабованості, гарантуючи, що кожного разу, коли масштаб системи збільшується вгору чи вниз, немає потреби перез'єднувати всі

ключі або вносити зміни до всіх серверів. Однак питання розподілу даних між серверами залишається відкритим. Тобто, може виникнути ситуація, коли розподіл серверів за хеш-кільцем нерівномірний, тобто розмір частин, за які відповідає кожен з серверів, не однаковий. Наприклад, ми почали з трьох серверів (сервер 0, сервер 1, сервер 2), які були більш-менш рівномірно розподілені вздовж кільця. Якщо один з серверів вийде з ладу, то навантаження, яке отримає наступний від нього сервер, буде вище (за умови, якщо всі дані, що надходять, мають рівномірний розподіл ключів). Насправді, це питання є значно складнішим, адже в більшості випадків дані не мають рівномірного розподілу. Таким чином, два вищеописані фактори в сукупності можуть призвести до ситуації, коли один з серверів отримує дуже високе навантаження (рисунок 3.22).

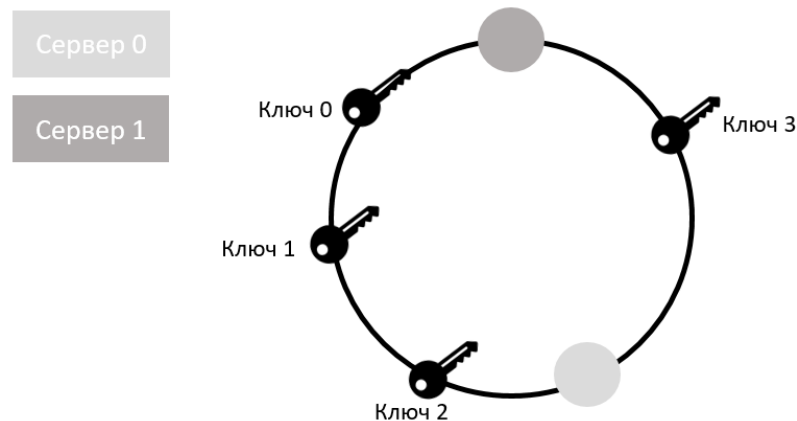


Рисунок 3.22 - Нервномірний розподіл ключів між серверами на хеш-кільці

Ця проблема може бути вирішена через введення кількох реплік або віртуальних вузлів для кожного сервера вздовж кільця. Наприклад, сервер 0 може мати дві репліки, розміщені в різних точках хеш-кільця (рисунок 3.23).

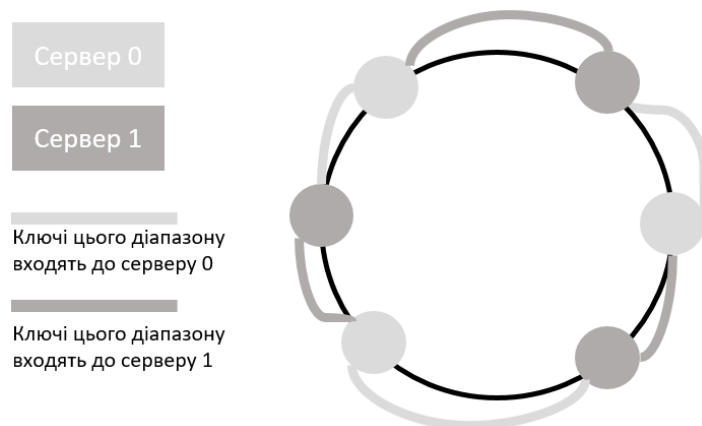


Рисунок 3.23 - Використання віртуальних вузлів для рівномірно розподілу ключів в хеш-кільці

По мірі збільшення кількості реплік або віртуальних вузлів у хеш-кільці розподіл ключів стає все більш рівномірним. У реальних системах кількість віртуальних вузлів/реплік дуже велика (більше 100). Якщо ми додамо по одній репліці серверів 0 і 1, то розподіл ключів буде рівномірним - 50% ключів надходять до серверу 0, 50% - до серверу 1 (рисунок 3.24).

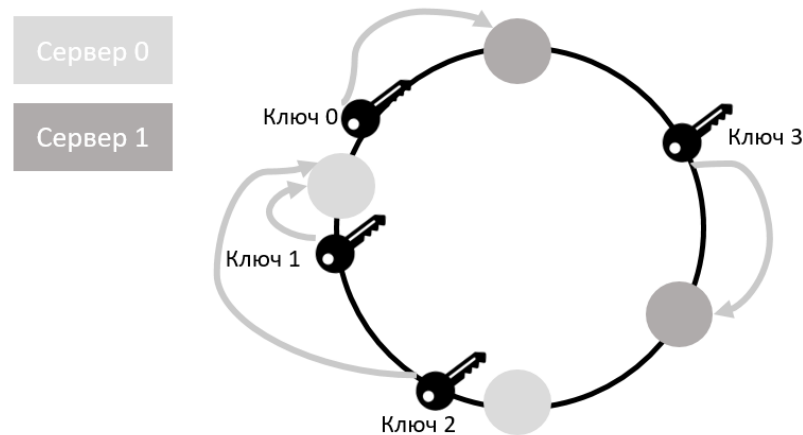


Рисунок 3.24 - Рівномірний розподіл ключів в хеш-кільці

Таким чином, консистентне хешування:

- а) робить можливим еластичне масштабування кластеру серверів баз даних /кешу;
- б) спрощує реплікацію та розподіл даних на серверах;
- в) забезпечує їх рівномірний розподіл, що прибирає «гарячі точки» (перевантажені сервери) завдяки опції розбиття даних;
- г) підвищує доступність системи в цілому.

3.3 База даних Apache Cassandra

В світі технологій баз даних існує два основних напрямки: SQL і NoSQL, реляційні та нереляційні бази даних. Відмінності між ними полягають в тому, як вони спроектовані, які типи даних підтримують та як зберігають інформацію. Реляційні бази даних зберігають структуровані дані, які зазвичай представляють об'єкти реального світу. Наприклад, відомості про людину, або про вміст кошика для товарів в магазині, згруповані в таблицях, формат яких заданий ще на етапі

проектування цих таблиць. Нереляційні баз даних влаштовані інакше і відрізняються способами запису даних (рисунок 3.25).

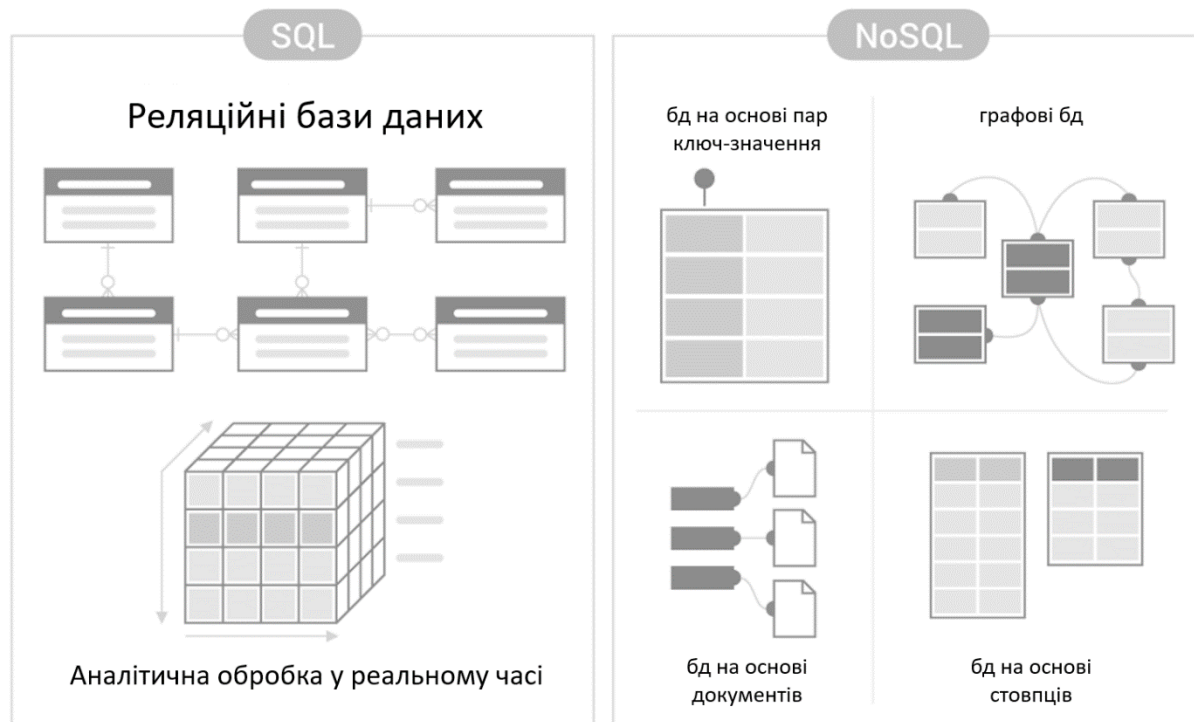


Рисунок 3.25 - Типи SQL та NoSQL баз даних

Наприклад, документно-орієнтовані нереляційні бази зберігають інформацію у вигляді ієрархічних структур даних, що можуть описувати об'єкти з довільним набором атрибутів. Те, що в реляційній базі даних розбивається на кілька взаємопов'язаних таблиць, в нереляційній може зберігатися у вигляді цілісної сутності [22].

Порівняльна характеристика SQL та NoSQL баз даних.

Структури даних і їх типи. SQL бази даних використовують строгі схеми даних, NoSQL бази даних допускають будь-який тип даних.

Запити. Незалежно від типу ліцензії, SQL бази даних в тій чи іншій мірі відповідають стандартам SQL, тому дані з них можна отримувати за допомогою мови SQL. NoSQL бази даних використовують специфічні способи запитів до даних, що робить NoSQL продукти специфічними у написанні запитів. NoSQL дані зберігаються багатьма способами (document-oriented, column-oriented, на основі графів або у якості key-value пар). Тому синтаксис змінюється від бази даних до бази даних [23].

Масштабованість. Обидва типи досить легко піддаються вертикальному масштабуванню (тобто, збільшенню системних ресурсів). NoSQL, як більш сучасний продукт, пропонує простіші способи горизонтального масштабування (тобто, створення кластера з декількох машин) [23].

Надійність. Коли справа доходить до збереження даних і гарантії виконання транзакцій SQL бази даних займають лідируючі позиції.

Підтримка. SQL бази даних дуже популярні і пропонують як платні, так і безкоштовні рішення. При виникненні проблем набагато простіше знайти відповідь, якщо справа стосується саме реляційних систем, особливо якщо рішення досить складне за своєю природою.

Зберігання та доступ до складних структур даних. З самого початку реляційні системи припускали роботу зі складними структурами, завдяки чому вони перевершують інші рішення по продуктивності.

Важливо розуміти, коли і який тип баз даних потрібно використовувати. Реляційні бази даних використовують тоді, коли:

а) є необхідність у відповідності бази даних вимогам ACID (Atomicity, Consistency, Isolation, Durability - атомарність, несуперечливість, ізолюваність, довговічність) [24];

б) дані структуровані, а їх структура не схильна до частих змін. Нереляційні бази даних використовують тоді, коли є підозри, що база даних може стати вузьким місцем проекту, що базується на роботі з великими обсягами інформації, зокрема коли:

а) передбачається зберігання великих обсягів неструктурованої інформації (NoSQL бази даних не накладають обмежень на типи даних, що зберігаються) [24];

б) передбачається використання хмарних обчислень і сховищ (хмарні сховища вимагають такого стану даних, щоб в разі необхідності їх можна було легко розподілити між кількома серверами для забезпечення масштабування);

в) передбачається швидка розробка (нереляційні бази даних не потребують того ж обсягу підготовчих дій, які зазвичай необхідні для реляційних баз)[24].

Таким чином, реляційні бази ставлять в свою основу 100% цілісність даних і статичність, а нереляційні - швидкість, гнучкість, масштабільність і динамічність.

Apache Cassandra є NoSQL базою даних. Касандра - це розподілена база даних від Apache, яка є масштабованою і призначена для управління дуже великою кількістю структурованих даних. Касандра створена для роботи з високомасштабованими і надійними сховищами величезних масивів даних [25]. Касандра призначена для обробки великого обсягу структурованих даних на товарних серверах. Касандра здатна обробляти таку кількість даних завдяки розподіленій архітектурі. Дані розміщуються на різних машинах з більш ніж одним коефіцієнтом реплікації, що забезпечує високу доступність і відсутність єдиної точки відмови. На рисунку 3.26 кола представляють вузли бази даних, а лінії між колами - розподілену архітектуру, в той час як клієнт надсилає дані у вузол.

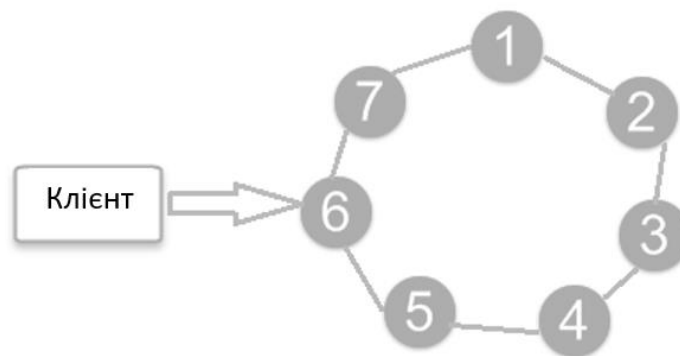


Рисунок 3.26 - Спрощена архітектура бази даних Касандра

Ключові особливості Касандри.

Відкритість. Касандра, незважаючи на потужність та надійність, є безкоштовною. Це проект з відкритим кодом від Apache. Завдяки відкритому коду вона породила величезну ком'юніті, де люди обговорюють свої запити та погляди. Крім того, існує можливість інтеграції Касандри з іншими проектами з відкритим кодом Apache, такими як Hadoop, Apache Pig, Apache Hive тощо. Це перша і важлива особливість Касандри.

Peer-to-peer архітектура. Деякі бази даних працюють з master-slave архітектурою в Касандрі, а деякі - з peer-to-peer архітектурою. У master-slave

архітектурі є основний блок, а решта спілкується з цим блоком. В той час як в peer-to-peer архітектурі кілька блоків спілкуються один з одним. Касандра дотримується peer-to-peer архітектури [25]. Саме тому Касандра не має жодної точки відмови і постійно доступна для критично важливих для бізнесу програм, які не можуть дозволити собі відмову. Також у жодному з центрів обробки даних не можна додати певну кількість вузлів до жодного кластеру. Касандра має надійну архітектуру з винятковими характеристиками.

Гнучка масштабованість. Одна з найбільших переваг. Касандра лінійно масштабується, тобто, збільшує пропускну здатність по мірі збільшення кількості вузлів у кластері [25]. Тому вона підтримує швидкий час реагування. За необхідності можна легко збільшити або зменшити масштаб кластеру. Касандра забезпечує гнучкість для додавання або видалення будь-якої кількості вузлів з кластеру без порушень її роботи. Тобто, не потрібно перезапустити кластер під час масштабування вгору чи вниз. Завдяки цьому Касандра має дуже високу пропускну здатність для найбільшої кількості вузлів. Крім того, під час масштабування спостерігається нульовий час простою або будь-яких пауз. Отже, пропускну здатність читання і запису збільшується одночасно, не затримуючись.

Гнучке зберігання даних. Касандра вміщує всі можливі формати даних, включаючи: структуровані, напівструктуровані та неструктуровані. Вона може динамічно вносити зміни до структури даних відповідно до потреб користувача.

Висока доступність і відмовостійкість. Реплікація даних - це зберігання даних у кількох місцях. Завдяки реплікації даних Касандра є високодоступною та стійкою до збоїв. Здебільшого, якщо один вузол кластеру виходить з ладу, дані доступні в інших вузлах. Тому можна отримати дані з цих вузлів. Кількість реплікацій визначається користувачем. Відповідно до цього числа, можна копіювати кожен рядок у кластері на основі його ключа (унікального ідентифікатора) [25]. Реплікація даних може здійснюватися в декількох центрах обробки даних. Очевидно, що це призводить до резервного копіювання та відновлення на високому рівні.

Висока працездатність. Касандра має одбу з найкращих показників порівняно з іншими NoSQL базами даних. Розробники хотіли використати

можливості багатьох багатоядерних машин, що стало основою розвитку бази даних. Касандра зарекомендувала себе як надзвичайно надійна, коли мова йде про великий набір даних. Тому вона використовується багатьма організаціями, які щодня обробляють величезну кількість даних.

Стовпчикова орієнтація. Модель даних Касандри орієнтована на стовпці, тобто, на зберігання даних таблиці за стовпцями, а не за рядками (рисунок 3.27).

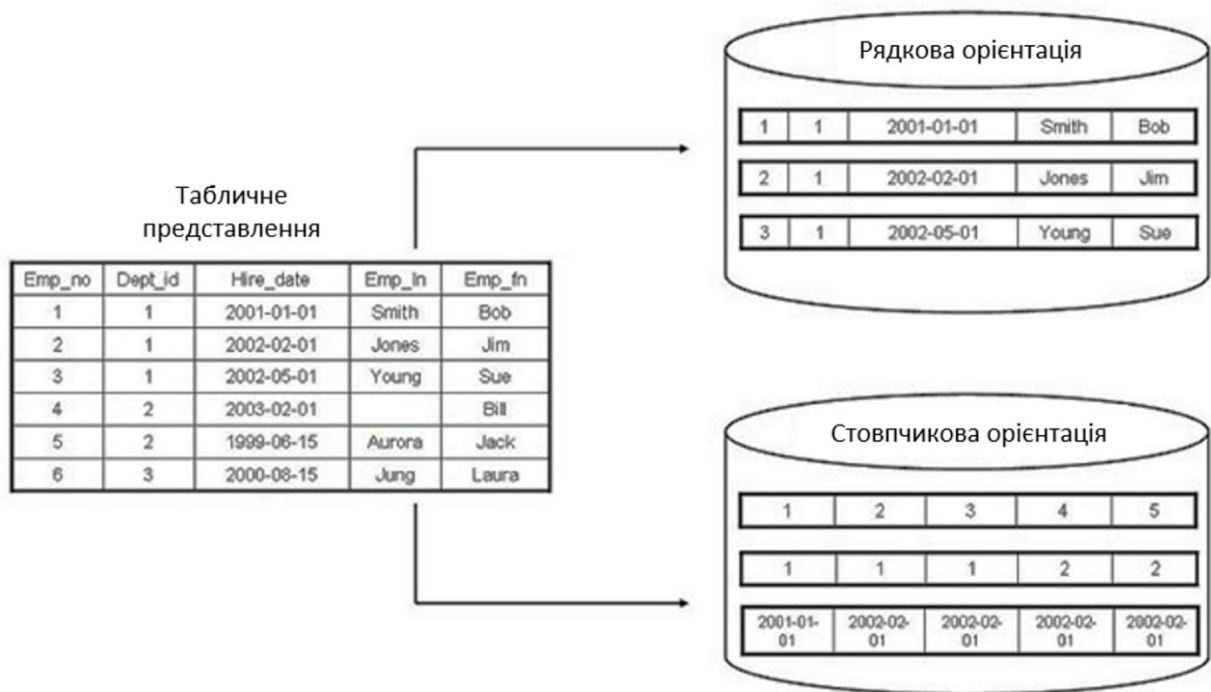


Рисунок 3.27 - Рядкова та стовпчикова орієнтація бази даних

Зберігаючи дані в стовпцях, а не в рядках, база даних може точніше отримувати доступ до даних для відповіді на запит замість сканування та відкидання небажаних даних у рядках.

Налаштовувана узгодженість. Узгодженість стосується того, як оновляти та синхронізувати рядок даних Касандри у всіх його репліках. В основному, існує два типи узгодженості в Касандрі: поточна узгодженість та сильна узгодженість [26]. Розробник може вибрати будь-який з них відповідно до його вимог. Як тільки кластер приймає запис, поточна узгодженість гарантує, що клієнт схвалює. В той же час, сильна узгодженість переконується, що будь-яке оновлення транслюється у всі вузли/машини, де конкретні дані є підходящими. Більше того, поєднання двох консистенцій також можливе.

Schema-free. Schema-free бази даних - це бази даних, в яких дані можуть зберігатися без попередньої структури [26]. В таких базах даних можна просто додавати записи без будь-якої попередньо визначеної структури. Крім того, можна групувати записи, які не мають однакової структури.

Головною метою Касандри є керування великими обсягами даних у кількох вузлах без жодної точки відмови. Касандра має розподілену систему peer-to-peer по своїх вузлах, а дані розподіляються між всіма вузлами кластеру. Усі вузли кластера відіграють однакову роль. Кожен вузол є незалежним і в той же час взаємопов'язаним з іншими вузлами. Кожен вузол кластера може приймати запити на читання та запис незалежно від того, де фактично розташовані дані в кластері. Коли вузол падає, запити на читання/запис можуть обслуговуватись іншими вузлами кластеру. В Касандрі один або кілька вузлів кластера виконують функції реплік для заданої частини даних. Якщо виявляється, що деякі з вузлів відповідають застарілим значенням, Касандра повертає клієнту останнє оновлене значення. Після повернення останнього значення Касандра виконує відновлення читання у фоновому режимі, щоб оновити застарілі значення. Касандра використовує реплікацію даних серед вузлів кластеру, щоб забезпечити відсутність єдиної точки відмови (рисунок 3.28).

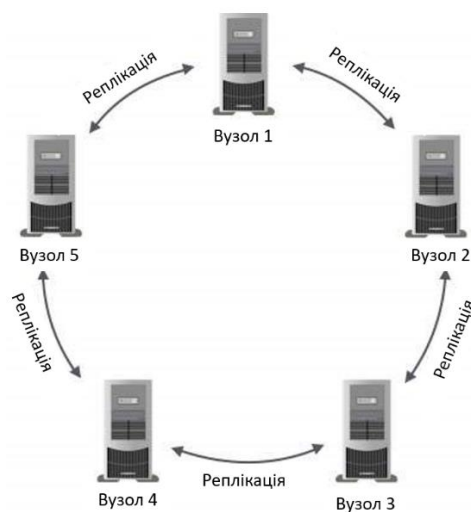


Рисунок 3.28 - реплікація даних в базі даних Касандра

Ключові поняття архітектури Касандри:

а) вузол - місце, де зберігаються дані; основна фундаментальна одиниця Касандри, в якій зберігаються дані (комп'ютер/сервер);

б) центр даних - сукупність взаємопов'язаних вузлів, є централізованим місцем для розміщення комп'ютерних та мережевих систем для задоволення потреб інформаційних технологій організації;

в) кластер - компонент, який містить один або кілька центрів обробки даних;

г) стійка - блок, який містить певну кількість серверів, які розміщені один поверх іншого;

д) журнал фіксації - механізм відновлення ланцюжків зміни даних у разі аварійних ситуацій у Касандрі;

е) MemTable (таблиця пам'яті) - тимчасове місце пам'яті, куди записуються дані під час оновлень або видалення;

ж) SSTable - файл диска, до якого дані видаляються з таблиці пам'яті, коли її вміст досягає порогового значення.

Модель даних Касандри істотно відрізняється від тієї, що зазвичай зустрічається в реляційних моделях. Касандра зберігає дані через кластери і розподіляється на декількох машинах, що працюють разом. Кожна машина працює як вузол і має власну репліку на випадок відмови машини. Вузли розташовані у форматі кластеру-кільця. Касандра keyspace (простір ключів) є контейнером для даних. Він містить багато атрибутів, основними з яких є:

- а) коефіцієнт реплікації;
- б) стратегія розміщення реплік;
- в) сімейство стовпців.

Коефіцієнт реплікації – це кількість машин у кластері, які отримують копії одних і тих же даних. Стратегія розміщення реплік – це стратегія розміщення реплік на кільці вузлів. Існує два типи стратегій: проста стратегія (для єдиного центру обробки даних), стратегія мережевої топології (для декількох центрів обробки даних) [24]. Сімейство стовпців – це простір ключів є контейнером для списку однієї чи більше сімей стовпців. Сімейство стовпців, у свою чергу, є контейнером набору рядків. Кожен рядок містить упорядковані стовпці. Сімейства стовпців представляють структуру даних (рисунок 3.29).

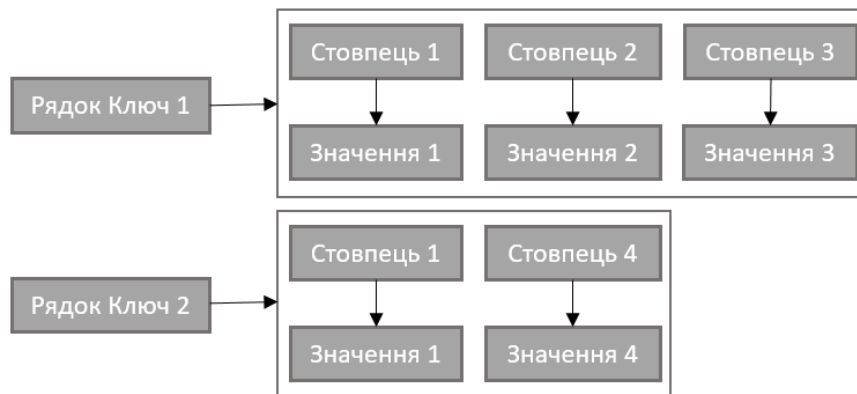


Рисунок 3.29 - Сімейство стовпців

Кожен простір ключів має щонайменше одне та часто багато сімейств стовпців (рисунок 3.30).

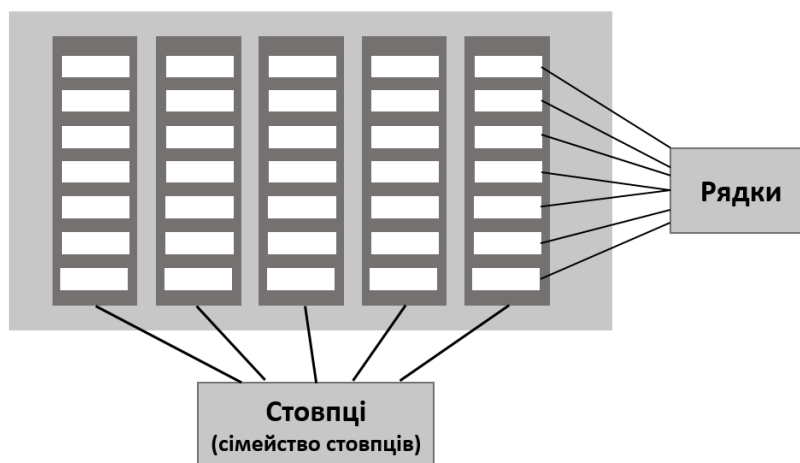


Рисунок 3.30 - Простір ключів

Рядки в кожному стовпчику - це сукупність багатьох стовпців. Стовпці є основною одиницею структури даних у Касандрі і складаються з трьох значень: ім'я ключа або стовпця, безпосередньо значення та часова мітка (рисунок 3.31).

Стовпець		
Ім'я : byte[]	Значення : byte[]	Час : clock[]

Рисунок 3.31 - Складові стовпця

Підсумовуючи, модель даних Касандри можна представити у вигляді шарів (рисунок 3.32).

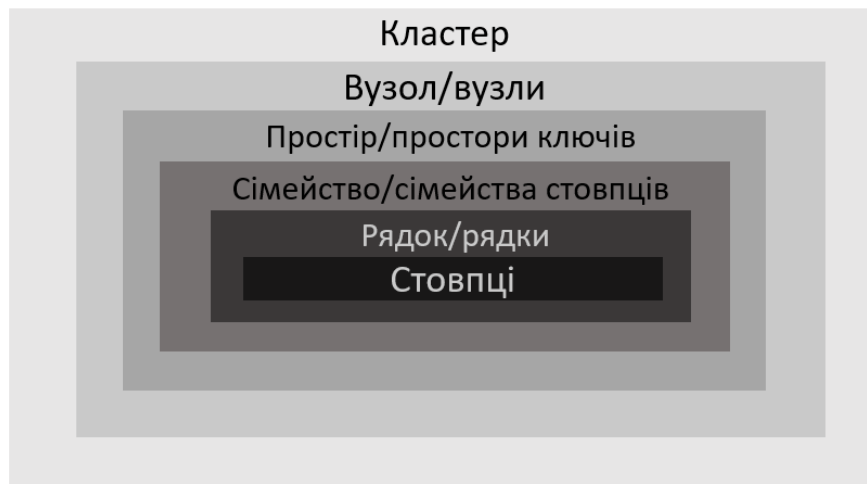


Рисунок 3.32 - Складові елементи кластеру

Вузол є другим шаром кластеру, що складається з систем/комп'ютерів або одиниць зберігання даних. Кожен кластер може містити багато вузлів, що з'єднані між собою і спільно обмінюються даними. Наступним шаром пам'яті є простір ключів. У вузлі є багато просторів ключів. Ці простори ключів, по суті, є найбільш зовнішнім блоком зберігання в системі. Вони містять основні дані. Далі простір ключів розподілений на сімейства стовпців. Ці сімейства стовпців мають різні області або заголовки, під якими поширюються дані. Сімейства стовпців також можна охарактеризувати як таблиці, що мають API. Наступним шаром у базі даних є рядки відповідно до сімейства стовпців. Рядки - це в основному класифікація, за якою поділяється сім'я стовпців. Ці класифікації, у свою чергу, створюють конкретні критерії розподілу для записів. Кінцевим шаром є стовпець, який розділений на різні заголовки, що містять основні дані щодо конкретного запису.

Основні переваги Касандри.

Швидкість запису. Касандра є одним з небагатьох сховищ даних, в якому запис відбувається швидше за читання.

Реплікація даних між центрами даних. Касандра здійснює реплікацію даних між різними центрами даних, причому, в необмеженій кількості.

Налаштовувана послідовність. Касандра забезпечує можливість вирішити, що станеться з даними при відмові одного чи декількох вузлів кластера.

Засновано на JVM (Java Virtual Machine - віртуальна машина Джава). Касандра написана на мові програмування Джава. Це означає, що вона може тісно інтегруватися з іншими програмами на базі JVM, включаючи безліч популярних інструментів для роботи з даними, таких як: Apache SOLR, Spark, Hadoop, Hive, PIG, Mahout та багато інших. Крім цього, JVM має величезну кількість підтримки та інструментів для усунення різних проблем, які можуть виникнути.

Мова запитів CQL (Cassandra Query Language). Це декларативна мова, яка дозволяє користувачам писати запити в Касандрі, використовуючи мову, що схожа на SQL [26].

Основні недоліки Касандри.

Відсутність ad-hoc (спеціальних) запитів. Спеціальні запити є вузькоспеціалізованими запитами, що створені для вирішення одноразових проблем[27]. Шар зберігання даних в Касандрі представлений системою зберігання даних у форматі «ключ-значення». Це означає, дані потрібно «модельовати» навколо запитів, які потенційно будуть виконуватись, а не навколо структури самих даних. Це може призвести до збереження одних і тих же даних різними способами задля можливості виконувати конкретні запити.

Непередбачувана продуктивність. Оскільки у Касандрі є багато різних асинхронних та фонових завдань, які не заплановані користувачем, продуктивність може бути непередбачуваною.

Засновано на JVM. Крім переваг, управління пам'яттю Касандри здійснюється мовою, а не програмою. При роботі з великими обсягами інформації необхідно налаштовувати JVM відповідно до конкретних потреб [27].

3.4 Архітектура, що горизонтально масштабується

Фронтенд буде побудовано за принципом SPA - Single Page Application, фронтенд та бекенд будуть розв'язані за допомогою API. Фронтенд буде спілкуватись з API через Load Balancer, який забезпечить можливість горизонтально масштабувати API. API буде створено з використанням мікросервісної архітектури. API буде вести комунікацію з базою даних. Приклад комунікації зображено на рисунку 3.33.

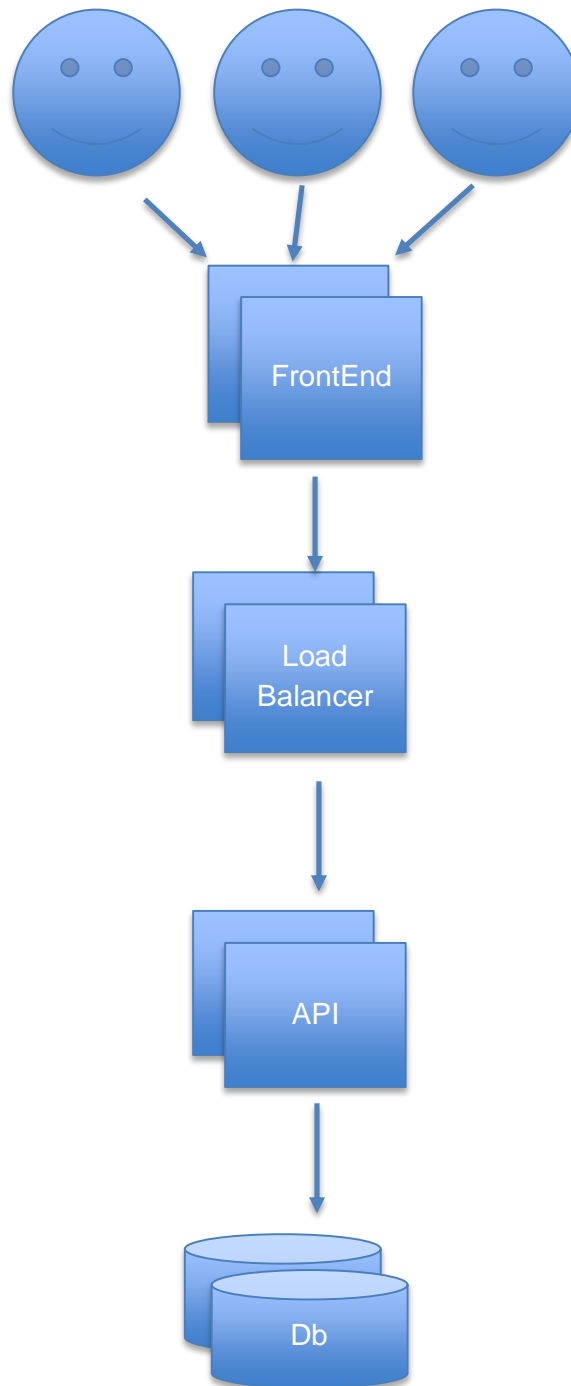


Рисунок 3.33 - Концептуальна архітектура

База даних повина також горизонтально масштабуватися для обслуговування зростаючого навантаження - для цього підійде Cassandra [дод. А, пл.3]. Так як ми будемо мати справу з розподіленими транзакціями (штрафи, винагороди), ми забезпечимо фінансову узгодженість через використання паттерну SAGA. Мікросервіси у шарі API можуть бути написані на будь-якій високорівневій мові програмування: C# / Java / Python. Загалом система повина розміщуватись у хмарному середовищі, яке надасть необхідну кількість ресурсів

для роботи системи за будь-якого навантаження. Наприклад, можна використати Google Cloud Platform. Модульна архітектура системи включає в себе декілька основних компонентів, які будуть наведені нижче.

Сервіс інтеграції з клієнтами - дозволяє підтримувати обмін даними з клієнтами системи, включаючи інформацію про «гаманець».

Сервіс користувачів - зберігає персональні дані користувача: унікальний ідентифікатор, ім'я, псевдонім (та ін.).

Сервіс контенту користувачів - зберігає інформацію про контент, що створено користувачем. Кожний вид контенту має унікальний ідентифікатор, що надає можливість пов'язати його з користувачем та (можливою) скаргою або апеляцією.

Сервіс скарг - веде скарги, що прив'язані до конкретного користувача. Користувач може мати різне відношення до скарги - бути тим, на кого створено скаргу; тим, хто сворив скаргу на іншого користувача; бути суддею під час судового етапу; бути суддею під час апеляції на прийняте рішення.

Сервіс фінансових транзакцій - дозволяє надавати винагороди та штрафувати користувачів системи. Я бачу реалізацію цього модуля за допомогою використання Blockchain, щоб забезпечити валідацію об'єму грошей системи.

3.5 Висновок

В розділі розглянуто можливі підходи для створення архітектури, що горизонтально масштабується. В основі горизонтально масштабованих систем лежить можливість горизонтального масштабування сховища даних. У випадку даної системи ця можливість забезпечується механізмом консистентного хешування. Саме цей механізм лежить в основі відомої NoSQL бази даних Apache Cassandra. Так як NoSQL база не забезпечує ACID, потрібно використати механізм SAGA для розподіленої транзакції. З урахуванням цих вимог та особливостей було створено можливу концептуальну архітектуру системи. Реалізація цієї архітектури планується у майбутніх дослідженнях.

4 РОЗРОБКА СТАРТАП-ПРОЕКТУ

4.1 Опис ідеї проекту

Ідея проекту полягає у створенні системи консенсусу, що зможе забезпечити реалізацію децентралізованого контролю якості користувацького контенту. Дана система представляє собою цілісний продукт, який є альтернативою існуючих служб підтримки соціальних медіа. Можливі напрямки застосування продукту (табл. 4.1):

- а) повна автоматизація контролю користувацького контенту;
- б) часткова автоматизація контролю користувацького контенту;
- в) забезпечення єдиного підходу для контролю якості контенту.

Повна автоматизація контролю користувацького контенту. Система забезпечує повну автоматизацію механізмів збору скарг, визначення суду присяжних, обробки рішень винесених членами суду присяжних, стягнення штрафів та надання винагород, подання та розгляду апеляцій. В кінцевому підсумку, не автоматизованим є лише процес прийняття рішення про прийнятність/неприйнятність конкретного контенту, однак цей етап реалізується самими користувачами в межах системи консенсусу. Така автоматизована система може стати повноцінною заміною стандартних служб підтримки, які сьогодні реалізують контроль якості контенту в соціальних мережах. Обираючи такий напрямок компанія приймає на себе ризики, що пов'язані з впровадженням нових технологій. Витрати на забезпечення роботи служб підтримки відсутні. Середня швидкість розгляду скарг суттєво збільшується.

Часткова автоматизація контролю користувацького контенту. Всі вищеописані процеси в системі так же автоматизовані, однак система може лише частково замінити службу підтримки (наприклад, для одної частини продукту чи проекту). Тобто, компанія може комбінувати старий і новий підходи до контролю контенту. Обираючи такий напрямок застосування системи компанія здійснює плавне впровадження технології в свої процеси, за рахунок чого мінімізує свої ризики. Витрати на забезпечення роботи служб підтримки зменшуються залежно від делегованого системі обсягу роботи. Крім того,

компанія може порівняти результати використання старого і нового підходів і визначити підходящість технології саме для її екосистеми. Середня швидкість розгляду скарг відносно збільшується, залежно від делегованого системі обсягу роботи.

Забезпечення єдиного підходу для контролю якості контенту. Розроблювана система може стати інструментом для уніфікації процесів контролю користувацького контенту в глобальному віртуальному просторі. Таке застосування технології здатне мінімізувати втручання держав у процес контролю користувацького контенту, підтримувати цілісність та неподільність віртуального інформаційного простору та сприяти міжнародній глобалізації.

Таблиця 4.1 - Опис ідеї стартап-проекту

<i>Зміст ідеї</i>	<i>Напрямки застосування</i>	<i>Вигоди для користувача</i>
Система консенсусу, що забезпечує механізм реалізації децентралізованого контролю якості користувацького контенту	Повна автоматизація контролю користувацького контенту (повна заміна діючих служб підтримки)	Відсутність витрат на забезпечення роботи служб підтримки; оптимізація процесів; суттєве збільшення швидкості розгляду скарг та підвищення пропускної здатності для обробки скарг
	Часткова автоматизація контролю користувацького контенту (часткова заміна діючих служб підтримки)	Відносне зменшення витрат на забезпечення роботи служб підтримки; оптимізація процесів; відносне збільшення швидкості розгляду скарг та підвищення пропускної здатності для обробки скарг; мінімальна ризиковість
	Забезпечення єдиного підходу для контролю якості користувацького контенту в глобальному віртуальному просторі	Мінімізація втручання держав в процес контролю користувацького контенту; сприяння глобалізації; підтримка відкритості та неподільності Інтернет простору; прозорість процесу контролю якості користувацького контенту

Техніко-економічні властивості та характеристики ідеї:

- а) легка масштабованість;
- б) низька ресурсозатратність впровадження;
- в) низька ресурсозатратність підтримки;
- г) висока гнучкість процесу оцінки контенту в контексті суб'єктів оцінки;
- д) висока гнучкість процесу оцінки контенту в контексті середовища оцінки;

е) середня ризикованість.

Легка масштабованість. Система представляє реалізацію сукупності правил що застосовуються на певній множині користувачів. Масштабування системи передбачає застосування цієї ж сукупності правил, але на більшій множині користувачів, за рахунок чого масштабування обмежується виключно технічними можливостями, необхідними для реалізації роботи системи з більшою кількістю користувачів.

Низька ресурсозатратність впровадження та підтримки. Вартість впровадження системи обмежується ресурсами, що використовуються для інтеграції нової технології в процеси компанії. Вартість використання системи для розгляду однієї скарги становить 1 цент, що в кінцевому підсумку є невисокою сумою для компанії будь-якого масштабу.

Висока гнучкість процесу оцінки контенту. По-перше, система забезпечує гнучкий процес оцінки контенту з точки зору суб'єктів, що його оцінюють. Кожен раз конкретний контент оцінюється різними, випадково обраними користувачами. Крім того, остаточне рішення про прийнятність/неприйнятність даного контенту утворюється в результаті агрегації рішень різних користувачів, що мінімізує суб'єктивізм та вплив людського фактору на остаточне рішення. По-друге, система забезпечує гнучкий процес оцінки контенту з точки зору середовища, в якому його оцінюють. Кожна одиниця контенту оцінюється випадково обраними користувачами, що є представниками глобального віртуального середовища. Таким чином, мінімізується вплив положень компаній або законодавства конкретних держав на остаточне рішення про прийнятність/неприйнятність контенту.

Середня ризикованість. Ідея є ризиковано зважаючи на те, що вона представляє кардинально новий підхід до контролю користувацького контенту.

На сьогоднішній день на ринку не існує технологій/концепцій, які можна було б розглядати в якості конкурентів системи консенсусу, однак існують товари-замінники – діючі служби підтримки компаній. Техніко-економічні властивості та характеристики концепції служб підтримки:

- а) обмежена масштабованість;
- б) середня ресурсозатратність впровадження;
- в) висока ресурсозатратність підтримки;
- г) низька гнучкість процесу оцінки контенту в контексті суб'єктів оцінки;
- д) низька гнучкість процесу оцінки контенту в контексті середовища оцінки;
- е) низька ризикованість.

Обмежена масштабованість. Здатність до масштабування служб підтримки обмежується не лише людським ресурсом, а і грошовим, адже з кожним новим працівником компанія витрачає частину ресурсів перш за все на його пошуки, а потім на ризики, що пов'язані з плинністю кадрів/необхідністю донавчання працівників/адаптацією працівників до процесів компанії тощо.

Висока ресурсозатратність впровадження. Як вже зазначилось вище, процес інтеграції нових працівників в службу підтримки досить ресурсозатратний, так як пов'язаний з рядом невід'ємних ризиків. Крім того, в більшості випадків новий працівник починає приносити реальну користь компанії не відразу. Ресурсозатратність підтримки роботи служби підтримки залежить від ряду факторів, але це в будь-якому випадку оплата праці, що встановлює відповідний рівень витрат.

Низька гнучкість процесу оцінки контенту. Процес прийняття рішень про прийнятність/неприйнятність контенту при роботі служб підтримки є низькою. По-першу, тому що кожен раз контент оцінюється постійною командою індивідумів. По-друге, тому що остаточне рішення про прийнятність/неприйнятність даного контенту утворюється в результаті агрегації поглядів однієї людини, що додає суб'єктивізму та впливу людського фактору до процесу оцінки контенту. По-третє, контент розглядається людьми, що представляють інтереси компанії, тому оцінка буде похідною від компанії та уряду деякої держави. Однак, рішення може бути адаптовано до законів конкретних країн у випадку надходження скарг від їх урядів.

Порівнюючи та підсумовуючи результати аналізу характеристик старого та нового підходів до регулювання контенту, система консенсусу має більшість переваг (табл. 4.2). Вона легше і з меншими затратами масштабується, вона вимагає менше ресурсів на впровадження та підтримку і передбачає більш гнучкий підхід до реалізації оцінки контенту. В той же час, як і будь-яка нова технологія вона є ризикованою і це варто враховувати.

Таблиця 4.2 - Визначення сильних, слабких та нейтральних характеристик ідеї проекту

№ n/ n	Техніко-економічні характеристики ідеї	(потенційні) товари/концепції конкурентів		W (слабка сторона)	N (нейтральн а сторона)	S (сильна сторона)
		Мій проект	Діючі служби підтримки			
1	Масштабованість	Обмежена технічними можливостями	Обмежена людськими ресурсами			+
2	Ресурсозатратність впровадження	Низька	Середня			+
3	Ресурсозатратність підтримки	Низька	Висока			
4	гнучкість процесу оцінки контенту в контексті суб'єктів оцінки	Висока	Низька			+
5	гнучкість/адаптивність процесу оцінки контенту в контексті середовища оцінки	Висока	Середня			+
6	Ризикованість	Середня	Низька	+		

4.2 Технологічний аудит ідей проекту

Система консенсусу є результатом роботи програмного коду, який не потребує залучення специфічних або важкодоступних технологій. Для написання коду потрібна високорівнева мова програмування. Для написання логіки системи потрібен ряд залежностей та бібліотек, які можна інтегрувати в код системи через використання відповідного фреймворку. Крім того, використання фреймворку значно спростить написання коду та його кінцеву архітектуру (табл. 4.3).

Таблиця 4.3 - Технологічна здійсненність ідеї проекту

№ n/n	Ідея проекту	Технології і реалізації	Наявність технологій	Доступність технологій
1	Система консенсусу для децентралізованого контролю якості контенту	Мова програмування C# та фреймворк .Net для програмування логіки системи	Технології наявні на ринку і є готовими інструментами для розробки системи	Технології доступні
2	Система консенсусу для децентралізованого контролю якості контенту	Мова програмування Java та фреймворк JavaSpring для програмування логіки системи	Технології наявні на ринку і є готовими інструментами для розробки системи	Технології доступні
Обрана технологія реалізації ідеї проекту: мова програмування C# та фреймворк .Net				

Технологічна реалізація проекту є можливою і не потребує залучення специфічних інструментів. В якості технологій, що будуть застосовані для розробки системи консенсусу для децентралізованого контролю якості контенту обрано мову програмування C# та фреймворк для програмування .Net. Обрані технології є доступними, легкими у використанні, розповсюдженими на ринку інформаційних технологій та мають розвинуту спільноту.

4.3 Аналіз ринкових можливостей запуску стартап-проекту

Використання соціальних медіа у всьому світі постійно зростає. Вони є найпопулярнішими онлайн-активностями, якими займаються користувачі мережі Інтернет. Станом на 2019 рік статистика соціальних медіа показує, що у світі існує 3,2 мільярди користувачів соціальних медіа, і це число лише зростає. Це становить приблизно 42% населення планети [31]. Крім того, очікується, що до 2021 року у світі буде понад 3 мільярди активних користувачів соціальних мереж. Порівняно з очікуваними в 2021 році 8 мільярдами населення можна очікувати, що у 2021 році частка користувачів, що активно використовують (а отже активно генерують та споживають контент) соціальні медіа, буде становити близько 40% [32]. Враховуючи всі вищенаведені факти, ринок соціальних медіа є великим та перспективним ринком, що постійно розвивається.

Згідно з інфографікою «Global social media research summary 2019», на сьогоднішній день ключовими гравцями на ринку соціальних медіа платформ за кількістю зареєстрованих та активних користувачів є Facebook та YouTube [30]. Динаміка ринку позитивна, адже він знаходиться у стані постійного зростання. Зростає не лише кількість користувачів, а і відсоток активних користувачів від загальної кількості. При цьому на ринку немає обмежень для входу нового продукту, адже він не має аналогів і може зайняти власну нішу на ринку соціальних медіа (табл. 4.4).

Таблиця 4.4 - Попередня характеристика потенційного ринку стартап-проекту

<i>№ п/п</i>	<i>Показники стану ринку (найменування)</i>	<i>Характеристика</i>
1	Кількість головних гравців, од	2 (Facebook, Youtube)
2	Загальний обсяг продаж, грн/ум.од	-
3	Динаміка ринку (якісна оцінка)	Зростає (соціальні медіа активно розвиваються)
4	Наявність обмежень для входу (вказати характер обмежень)	Відсутні, система консенсусу для децентралізованого контролю якості контенту є унікальною в своєму роді
5	Специфічні вимоги до стандартизації та сертифікації	-
6	Середня норма рентабельності в галузі (або по ринку), %	-

Таким чином, ринок є привабливим для входження описаної технології за двома причинами:

а) ринок соціальних медіа активно зростає. Станом на 2018 рік збільшується як кількість активних користувачів популярних платформ, так і обсяг контенту, що генерується користувачами соціальних медіа [28, 29];

б) на ринку не існує аналогів і прямих конкурентів даної технології.

Потенційними групами клієнтів для розроблюваної технології є:

а) великі компанії, для яких першочергова підтримка іміджу;

б) середні та великі компанії, для яких першочергова оптимізація процесів та витрат.

Великі компанії, для яких ключовими факторами є конкурентоспроможність, підтримка відповідного іміджу, задоволеність користувачів та залучення новітніх технологій як для покращення досвіду користувачів та підтримки конкурентоспроможності, так і для оптимізації процесів. Такі компанії сфокусовані на результатах роботи системи; якісних показниках, що свідчать про ступінь задоволеності користувачів від користування їх сервісом; кількісних показниках, що свідчать про обсяги користувачів/активних користувачів тощо. Для таких компаній вирішальним фактором при прийнятті рішення про використання технології буде порівняння ступеня задоволеності користувачів сервісу до та після впровадження системи. Основні вимоги до системи: простота в підтримці продукту, масштабованість продукту, наявність підтримки продукції з боку компанії-постачальника, репрезентативність та числове вираження кількісних та якісних результатів роботи системи. Основні вимоги до компанії постачальника: прозорість процесів, гарна репутація, надійність, постійна доступність для зворотнього зв'язку.

Середні компанії, для яких ключовими є оптимізація внутрішніх процесів та мінімізація витрат. Такі компанії відчують себе комфортно в їх середовищі і з обережністю ставляться до впровадження нових підходів. Однак, вони готові впроваджувати нові технології у разі, коли їх практична значущість переважає потенційні ризики. Такі компанії будуть зацікавлені в кількісних показниках роботи системи та її пропускній здатності. Для таких компаній вирішальним фактором при прийнятті рішення про використання технології буде порівняння витрат на реалізацію контролю користувацького контенту до та після впровадження системи. Основні вимоги до системи: простота в підтримці продукту, невисока вартість підтримки системи порівняно з застосуванням старих підходів, репрезентативність та числове вираження кількісних результатів роботи системи. Основні вимоги до компанії постачальника: прозорість процесів, надійність (табл. 4.5).

Таблиця 4.5 - Характеристика потенційних клієнтів стартап-проекту

<i>№ n/n</i>	<i>Потреба, що формує ринок</i>	<i>Цільова аудиторія (цільові сегменти ринку)</i>	<i>Відмінності у поведінці різних потенційних цільових груп клієнтів</i>	<i>Вимоги споживачів до товару</i>
1	Потреба в якісному контенті	Великі компанії, що турбуються про підтримку свого іміджу	Фокус на результатах роботи системи. Порівняння ступеня задоволеності користувачів до та після використання системи.	- до продукції: простота підтримки, масштабованість, наявність підтримки з боку компанії-постачальника, репрезентативність кількісних та якісних результатів роботи системи; - до компанії постачальника: прозорість процесів, гарна репутація, надійність, зворотній зв'язок;
2	Потреба в оптиміза ції процесів	Середні та великі компанії, що турбуються про оптимізацію процесів та витрат	Фокус на пропускній здатності роботи системи та процесах. Порівняння витрат до та після використання системи.	- до продукції: простота в підтримки, невисока вартість підтримки, репрезентативність кількісних результатів роботи системи; - до компанії постачальника: прозорість процесів, надійність;

Основні фактори, що можуть перешкодити ринковому впровадженню продукту:

- а) технологічна новизна продукту;
- б) відсутність репутації постачальника послуги.

Технологічна новизна продукту. Зважаючи на те, що технологія не має аналогів, а її концепція є унікальною в своєму роді, існує висока вірогідність недовіри з боку компаній та самих користувачів, що будуть використовувати дану систему. Тому варто забезпечити прозорість системи, налаштувати канали для зворотнього зв'язку і брати активну участь у його наданні. Зворотній зв'язок при правильному використанні може стати інструментом для підвищення довіри з боку користувачів системи, тому варто з самого початку усвідомлювати цінність фідбеку та забезпечити середовище для його руху.

Відсутність репутації постачальника послуги. Репутація постачальника продукту є вирішальним фактором для компаній, особливо коли мова йде про

продукт, якому буде делегуватись частина процесів компанії. Яким би якісним не був продукт, споживачі повинні хотіти його спробувати. Тому важливо попередньо залучитись підтримкою впливових суб'єктів ринку та забезпечити прозорість процесів та самої системи, щоб мінімізувати недовіру до новітнього продукту. Також важливо зробити процес впровадження технологій для компаній-споживачів максимально плавним та комфортним (наприклад, через надання безкоштовного тестового періоду). Особливу увагу потрібно приділити збору даних про роботу системи та формуванню відповідної звітності не лише для реклами продукту, а і для подальшого розвитку системи та моніторингу зміни показників (табл. 4.6).

Таблиця 4.6 - Фактори загроз

<i>№ n/n</i>	<i>Фактор</i>	<i>Зміст загрози</i>	<i>Можлива реакція компанії</i>
1	Технологічна новизна продукту	Унікальність та новизна концепції продукту, що створює ризики для потенційного споживача	Забезпечення прозорості у роботі системи, відкритість до діалогу з користувачами, активне використання зворотнього зв'язку
2	Відсутність репутації постачальника послуги	Репутація постачальника послуг є вирішальним фактором для компаній коли мова йде про нові технології	Залучення підтримки компаній, демонстрація результатів роботи системи, забезпечення тестової версії

Основний фактор, що може сприяти ринковому впровадженню продукту – унікальність продукту та новизна підходу, який він пропонує. Це створює можливість до захоплення частки існуючого ринку або створення та подальший розвиток ринкової ніші, що займається питаннями реалізації контролю якості контенту в глобальному віртуальному просторі. Тову важливо з самого початку створити відповідне бачення та концепцію продукту і активно використовувати їх у позиціонуванні продукту на ринку. Також, враховуючи можливість зростання частки ринку/ринкової ніші, що займає продукт, важливо від самого початку закласти в нього гнучкість і постійно адаптувати відповідно до потреб різних цільових аудиторій (табл. 4.7).

Таблиця 4.7 - Фактори можливостей

№ n/n	Фактор	Зміст можливості	Можлива реакція компанії
1	Унікальність продукту та новизна підходу	Захоплення частини ринку, розвиток ринкової ніші	Просування відповідної концепції продукту, забезпечення його адаптивності та гнучкості для задоволення потреб різних сегментів віртуального простору

Особливості конкурентного середовища:

- а) чиста монополія - продукт не має аналогів та прямих конкурентів, на ринку буде існувати лише одна компанія-постачальник продукту такого типу;
- б) світовий рівень кокурентної боротьби - конкуренція в міжнародному просторі; постачання продукту компаніям, що функціонують в глобальному віртуальному просторі;
- в) внутрішньогалузевість – продукт функціонуватиме в сфері інформаційних технологій і потенційно може сформувати окрему підгалузь – регулювання мережевого контенту;
- г) товарно-родова конкуренція - конкуренція між різними видами продуктів, що можуть виконувати схожі функції, тобто, конкуренція з боку продуктів-субститутів (замінників). Єдиний субститут розроблюваної системи – діючі служби підтримки в компаніях;
- д) нецінові конкурентні переваги - конкурентні переваги розроблюваної системи полягають у новизні підходу, на якому вона базується та кількісних/якісних перевагах нового підходу порівняно з існуючими підходами;
- е) не марочна інтенсивність - на ринку є лише один продукт-замінник - служби підтримки компаній.

Враховуючи специфіку конкурентного середовища продукту, для зайняття позиції на ринку або створення власної ніші потрібно розробити ряд заходів для створення сприятливого іміджу продукту та залучитись підтримкою компаній через співпрацю. Для закріплення та підтримки позиції на ринку та мінімізації вірогідності появи конкурентів в майбутньому необхідно забезпечити надійність та ефективність системи та здійснювати постійний моніторинг показників її

роботи і змінювати систему відповідно до потреб споживачів. Для витіснення продуктів-замінників потрібно демонструвати кількісні та якісні переваги продукту порівняно з результатами роботи діючих служб підтримки компаній (табл. 4.8).

Таблиця 4.8 - Ступеневий аналіз конкуренції на ринку

<i>Особливості конкурентного середовища</i>	<i>В чому проявляється дана характеристика</i>	<i>Вплив на діяльність підприємства (можливі дії компанії, щоб бути конкурентоспроможною)</i>
1. За типом конкуренції - чиста монополія	Відсутність аналогів на ринку	Постійний моніторинг показників ефективності та надійності
2. За рівнем конкурентної боротьби - світовий	Конкуренція в глобальному віртуальному просторі	Створення сприятливого іміджу та подальша підтримка репутації
3. За галузевою ознакою - внутрішньогалузева	Конкуренція в галузі інформаційних технологій	Захоплення та розвиток ніші
4. Конкуренція за видами товарів: - товарно-родова	Конкуренція з діючими службами підтримки компаній	Демонстрація кількісних та якісних переваг продукту
5. За характером конкурентних переваг - нецінова	Конкурентні переваги у новизні підходу	Демонстрація кількісних та якісних переваг продукту
6. За інтенсивністю - не марочна	Єдиний продукт-замінник – служби підтримки компаній	Демонстрація кількісних та якісних переваг продукту

На ринку немає прямих та потенційних конкурентів розроблюваного продукту. Постачальників продукт також не має, принаймні, на перших етапах його розвитку. Всі технології, що необхідні для його розробки та підтримки, знаходяться у вільному доступі. Головним фактором сили компаній-споживачів є їх служби підтримки. Адже вони вже мають готове та перевірене часом рішення, яке планується замінити новим рішенням, закладеним в основу розроблюваного продукту. З цієї точки зору компанії не мають потреби у розроблюваному продукті і можуть задовольнити потребу в інструменті для реалізації користувацького контролю своїми силами, без залучення сторонніх продуктів. Фактором загрози з боку діючих служб підтримки компаній є той наслідок того факту, що служби підтримки є частинами компаній. Компанії з

більшою вірогідністю будуть підтримувати існуючі рішення ніж впроваджувати нові (особливо коли мова йде про кардинально нові підходи).

Клієнти диктують умови роботи на ринку, зокрема: система має бути надійною та стійкою до непередбачуваних ситуацій; система має бути прозорою та гнучкою для того, щоб компанії могли впроваджувати її в свої процеси. Крім того, клієнти хочуть мінімізації ризиків для них у випадку впровадження даного продукту. Враховуючи специфіку роботи служб підтримки компаній, існує необхідність співпраці та конкурування з діючими продуктами (так як компанії будуть впроваджувати нову технологію поступово і не стануть одразу переходити на повністю автоматизований контроль контенту). Тому потрібно полегшити вхід технології на ринок і забезпечити підтримку конкурентоспроможності на рівні з існуючими рішеннями через забезпечення гнучкості системи. Тобто, система має бути достатньо гнучкою, щоб підлаштовуватись під специфічні потреби клієнтів і достатньо масштабованою, щоб працювати з різними обсягами користувачів залежно від потреб компанії-замовника. Так, на етапі введення нової технології (системи консенсусу) компанія може використовувати її лише частково, паралельно з використанням старої технології (діючої служби підтримки) для порівняння результатів старого і нового підходів (табл. 4.9).

Таблиця 4.9 - Аналіз конкуренції в галузі за М. Портером

<i>Складові аналізу</i>	<i>Прямі конкуренти в галузі</i>	<i>Потенційні конкуренти</i>	<i>Постачальник</i>	<i>Компанії соціальних медіа</i>	<i>Діючі служби підтримки компаній</i>
	-	-	-	Компанії володіють готовими рішеннями	Діючі служби підтримки є частинами компаній та не несуть ризиків
<i>Висновки</i>	-	-	-	Клієнти диктують умови роботи на ринку та вимагають від продукту надійності, мінімальної ризикованості та прозорості	Необхідність забезпечення достатнього рівня гнучкості, щоб співпрацювати з службами підтримки; необхідність забезпечення достатнього рівня конкурентоспроможності для поступового витіснення служб підтримки

Враховуючи той факт, що компанії з більшою вірогідністю будуть підтримувати існуючі рішення, ніж тестувати та впроваджувати нові (навіть якщо вартість підтримки існуючих рішень більша ніж потенційна вартість підтримки нових рішень). Тому на першому етапі важливо не стільки забезпечити конкурентоспроможність з товаром-замінником, скільки співпрацю з ним. Тобто, потрібно мінімізувати ризики для компаній, що будуть робити спроби переходу на нову технологію і зробити цей процес максимально плавним для них. Тобто, зробити систему досить гнучкою, щоб у разі необхідності вона могла працювати паралельно з діючою службою підтримки компанії/інтегруватись з нею/тощо. На другому етапі потрібно забезпечити наочність результатів роботи системи і перевагу кількісних/якісних результатів роботи нової системи порівняно з кількісними/якісними результатами роботи діючої служби підтримки. Тобто, показати сильні сторони застосування нового підходу порівняно із застосуванням старого підходу.

Враховуючи аналіз конкуренції, характеристики ідеї проекту та вимоги споживачів до продукту, можна виділити наступні фактори його конкурентоспроможності:

- а) масштабованість;
- б) вартість;
- в) пропускна здатність;
- г) якість роботи.

Продукт володіє здатністю до легкого масштабування. Фактор маштабованості є особливо важливим для великих компаній та для компаній, що активно розвиваються. Масштабованість служб підтримки обмежена чисельністю людських ресурсів компанії та потребує значних ресурсних витрат, в той час як масштабованість системи консенсусу обмежена лише технічними можливостями і майже не потребує ресурсних витрат. Процес масштабування системи представляє однотипний набір дій і в цілому є набагато легшим порівняно з процесом масштабування процесів та людей, який передбачає цілі сукупності підпроцесів.

Продукт володіє низькою вартістю підтримки. Вартість підтримки автоматизованої системи набагато нижча порівняно з вартістю організації та оплати праці людського ресурсу.

Продукт володіє високою пропускнуою здатністю та високою швидкістю опрацювання заявок. Пропускна здатність (кількість заявок, що можуть бути розглянуті за деяку одиницю часу) у автоматизованої системи в рази більша ніж у служби підтримки, що обмежується можливостями людського організму. Автоматизована система розгляне конкретну заявку в рази швидше порівняно зі службою підтримки.

Продукт володіє більшою потенційною якістю результатів роботи. Незважаючи на той факт, що система консенсусу автоматизована, остаточне рішення щодо прийнятності/неприйнятності контенту все одно приймається групою людей, що забезпечує конкурентоспроможність системи порівняно зі службами підтримки, де остаточне рішення приймається групою фахівців (табл. 4.10).

Таблиця 4.10 - Обґрунтування факторів конкурентоспроможності

<i>№ п/ п</i>	<i>Фактор конкурентоспро- можності</i>	<i>Обґрунтування (наведення чинників, що роблять фактор для порівняння конкурентних проектів значущим)</i>
1	Масштабованість	Масштабованість служб підтримки обмежена чисельністю людських ресурсів компанії, а масштабованість системи консенсусу обмежена лише технічними можливостями
2	Вартість	Низька вартість підтримки автоматизованої системи порівняно з вартістю підтримки служб підтримки
3	Пропускна здатність	Пропускна здатність автоматизованої системи в рази більша ніж у служби підтримки
4	Швидкість	Автоматизована система розгляне конкретну заявку в рази швидше порівняно зі службою підтримки.
5	Якість роботи	Рішення в системі приймаються групою людей, тоді як рішення в службі підтримки приймаються конкретними працівниками поодиночі

Сильними сторонами проекту є легка масштабованість, низька вартість впровадження та підтримки, висока швидкість обробки заявок, висока пропускна здатність системи та якість її роботи. Слабкими сторонами проекту є

певна непередбачуваність результатів роботи системи, що викликана новизною технології та відсутність репутації постачальника-продукту. Однак, враховуючи низьку вартість впровадження системи та можливість використання її у тестовому режимі, вплив даних факторів можна мінімізувати (табл. 4.11).

Таблиця 4.11 - Порівняльний аналіз сильних та слабких сторін «назва проекту»

№ n/n	Фактор конкурентоспроможності	Бали 1-20	Рейтинг діючих служб підтримки у порівнянні з системою консенсусу						
			-3	-2	-1	0	1	2	3
1	Масштабованість	20	+						
2	Вартість	20				+			
3	Пропускна здатність	20		+					
4	Швидкість	20	+						
5	Якість роботи	17							+
6	Передбачуваність результатів	10							+
7	Репутація	1							+

Враховуючи наведені сильні сторони проекту, існують наступні ринкові можливості:

а) залучення підтримки великих компаній (через надання їм безкоштовної пробної версії продукту для тестування та можливого подальшого впровадження);

б) фандрайзингова діяльність, залучення зацікавлених сторін через представлення пітч-деку для першого раунду залучення інвестицій;

в) адаптація системи до такого стану, коли вона зможе використовуватись в тестовому режимі/інтегруватись в процес роботи діючих служб підтримки задля мотивування компаній до переходу на нову технологію).

Враховуючи наведені слабкі сторони проекту, існують наступні ринкові загрози:

а) нижчий ступінь задоволеності користувачів організацією процесу контролю контенту порівняно з старим підходом до організації контролю контенту;

б) недовіра до технології з боку компаній-клієнтів та самих користувачів (табл. 4.12).

Таблиця 4.12 - SWOT- аналіз стартап-проекту

<p><i>Сильні сторони</i></p> <ul style="list-style-type: none"> – Масштабованість – Вартість – Пропускна здатність – Швидкість – Якість роботи 	<p><i>Слабкі сторони</i></p> <ul style="list-style-type: none"> – Певна непередбачуваність результатів – Відсутність репутації компаній-постачальника
<p><i>Можливості</i></p> <ul style="list-style-type: none"> – Залучення підтримки компаній – Фандрайзингова діяльність – Адаптація системи до можливості використання в тестовому режимі/інтеграції в процеси компаній 	<p><i>Загрози</i></p> <ul style="list-style-type: none"> – Нижчий ступінь задоволеності користувачів організацією процесу контролю контенту – Недовіра до технології

Враховуючи результати проведеного SWOT-аналізу, необхідно розробити альтернативи впровадження проекту, які змогли б мінімізувати загрози та застосувати сильні сторони для використання можливостей проекту. Першою альтернативою є адаптація системи для можливості попереднього тестування та/або поступового впровадження компаніями-замовниками задля збільшення ступеню зацікавленості компаній-замовників та мінімізації ризиків, які беруть на себе компанії при впровадженні нової технології. Така альтернатива мінімізує можливість виникнення недовіри до технології та дозволить порівняти ступінь задоволеності користувачів від застосування старого та нового підходів до контролю контенту. Крім того, дана альтернатива може посприяти залученню підтримки компаній. Другою альтернативою є фандрайзингова діяльність, що дозволить залучитись підтримкою інвесторів та вкласти кошти в рекламу продукту в першу чергу для інформування про продукт, а потім для підвищення впізнаваності, закріплення репутації продукту та мінімізації недовіри до технології. Дана альтернатива також може посприяти залученню підтримки компаній (табл. 4.13).

Таблиця 4.13 - Альтернативи ринкового впровадження стартап-проекту

№ n/n	Альтернатива (орієнтовний комплекс заходів) ринкової поведінки	Ймовірність отримання ресурсів	Строки реалізації
1	Адаптація системи для можливості попереднього тестування та/або поступового впровадження компаніями	Висока	6-12 місяців
3	Фандрайзингова діяльність	Середня	3-6 місяців

Серед перелічених альтернатив та, для якої отримання ресурсів є більш простим та ймовірним - перша; та, для якої строки реалізації є більш стислими - друга і третя. В кінцевому підсумку, обрана альтернатива - комбінування першої та другої альтернатив. Тобто, залучення підтримкою компаній паралельно з наданням цим компаніям тестових версій продукту для безпосередньої апробації ними системи та винесення подальшого рішення про співпрацю.

4.4 Розробка ринкової стратегії проекту

Основні цільові групи потенційних споживачів продукту:

- а) великі компанії, які турбуються про підтримку свого іміджу;
- б) середні та великі компанії, які турбуються про оптимізацію процесів та ресурсних затрат.

Великі компанії, які турбуються про підтримку свого іміджу, володіють високою готовністю до прийняття нового продукту, так як розуміють необхідність та практичну значущість впровадження нових підходів/технологій або вимушені активно конкурувати з іншими компаніями, що так чи інакше передбачає впровадження нових технологій. Орієнтовний попит в межах даної цільової групи високий. Компаній-представників даної цільової групи не так багато, однак більшість з них буде готова як мінімум спробувати нову технологію і, як максимум, прийняти її залежно від результатів апробації. Конкуренція в даному цільовому сегменті представлена лише продуктами-замінниками – службами підтримки. Простота входу у сегмент середня так як, незважаючи на відсутність прямих конкурентів, доведеться конкурувати з існуючими службами підтримки компаній.

Середні та великі компанії, які турбуються про оптимізацію процесів, володіють середньої готовністю до прийняття нового продукту, так як з одного боку, розуміють потенційні переваги від впровадження нових технологій, однак з іншого, розуміють ризики, пов'язані з впровадженням цих технологій. Для них важлива чітка перевага застосування нових підходів у порівнянні зі старими і в першу чергу вони оцінюють ризики. Орієнтовний попит в межах даної цільової групи середній. Конкуренція в даному цільовому сегменті представлена лише продуктами-замінниками – службами підтримки. Простота входу у сегмент

середня так як, незважаючи на відсутність прямих конкурентів, доведеться конкурувати з існуючими службами підтримки компаній (табл. 4.14).

Таблиця 4.14 - Вибір цільових груп потенційних споживачів

<i>№ п/п</i>	<i>Опис профілю цільової групи потенційних клієнтів</i>	<i>Готовність споживачів сприйняти продукт</i>	<i>Орієнтовн ий попит в межах цільової групи (сегменту)</i>	<i>Інтенсивні сть конкуренції в сегменті</i>	<i>Просто та входу у сегмент</i>
1	Великі компанії, які турбуються про підтримку свого іміджу	Висока	Високий	Низька	Середня
2	Середні та великі компанії, які турбуються про оптимізацію процесів та ресурсних затрат	Середня	Середній	Низька	Середня
Які цільові групи обрано: обидві					

За результатами аналізу обрано обидві цільові групи, тобто, передбачається робота з кількома сегментами аудиторії і розробка для них окремих програм ринкового впливу (стратегія диференційованого маркетингу).

Для роботи в обраних сегментах ринку сформувано базову стратегію розвитку, що передбачає адаптацію системи для можливості попереднього тестування та/або поступового впровадження компаніями-замовниками. По-перше, така стратегія дозволяє збільшити ступінь зацікавленості компаній-замовників та мінімізувати ризики для них, що в кінцевому підсумку збільшить кількість споживачів продукту. По-друге, така стратегія дозволяє залучитись підтримкою компаній. Стратегією охоплення ринку є цільовий маркетинг. Тобто, розмежування сегментів ринку, вибір одного або декількох цільових сегментів і подальша розробка прийомів маркетингу для кожного з обраних цільових сегментів окремо, цілеспрямовано, з урахуванням специфічних характеристик та вимог представників кожної цільової групи. Ключовою конкурентоспроможною позицією продукту відповідно до обраної альтернативи є надання товару важливих з точки зору споживача конкурентоспроможних властивостей, зокрема: адаптивність та гнучкість системи відповідно до потреб конкретних компаній-замовників, забезпечення тестової версії системи та/або версії системи, що може у разі необхідності інтегруватись з діючими службами підтримки компаній. Базовою стратегією розвитку обрано стратегію

диференціації - виготовлення особливого продукту, який являє собою варіант модифікації стандартного продукту. Наприклад, можна розробляти різні версії системи консенсусу, які будуть працювати з різними типами контенту/користувачів тощо (табл. 4.15).

Таблиця 4.15 - Визначення базової стратегії розвитку

<i>№ n/n</i>	<i>Обрана альтернатива розвитку проекту</i>	<i>Стратегія охоплення ринку</i>	<i>Ключові конкурентоспроможні позиції відповідно до обраної альтернативи</i>	<i>Базова стратегія розвитку</i>
1	Адаптація системи для можливості попереднього тестування та/або поступового впровадження компаніями	Цільовий маркетинг	Надання товару важливих з точки зору споживача та конкурентоспроможних властивостей	Стратегія диференціації

В якості стратегії конкурентної поведінки стартапу обрано стратегію виклику лідера. Така стратегія дозволить стати лідером ринку у разі застосування прийому флангової атаки (атака на слабкі сторони продукту-замінника - важка та ресурсозатратна масштабованість, ресурсозалежність, людський фактор та суб'єктивізм при оцінці контенту тощо). Незважаючи на те, що продукт є першопрохідцем і не має аналогів на ринку, він не може мати власних споживачів, так як всі потенційні споживачі продукту вже «зайняті». Вони вже використовують продукт-замінник (служби підтримки) і не мають альтернатив. Тому продукт буде забирати існуючих споживачів у конкурентів (споживачі - компанії, конкуренти - служби підтримки компаній, нова система буде приходити на зміну старим службам підтримки компаній). Новий продукт практично не копіює характеристики конкурентів. Єдина спільна риса між системою консенсусу та службою підтримки - процес створення скарг на контент користувачами (табл. 4.16).

Таблиця 4.16 - Визначення базової стратегії конкурентної поведінки

<i>№ n/ n</i>	<i>Чи є проект «першопрохід цем» на ринку?</i>	<i>Чи буде компанія шукати нових споживачів, або забирати існуючих у конкурентів?</i>	<i>Чи буде компанія копіювати основні характеристики товару конкурента, і які?</i>	<i>Стратегія конкурентної поведінки</i>
1	Система не має аналогів на ринку	Забирати існуючих у конкурентів	Процес створення скарг на контент користувачами	Стратегія виклику лідера

Стратегія позиціонування продукту передбачає розробку відповідної концепції та бачення, що направлені на формування та закріплення конкурентної позиції на ринку. Серед вимог, що висувуються споживачами до продукту: гнучкість; надійність; масштабованість; низька вартість підтримки порівняно з існуючими рішеннями; здатність замінити групу фахівців, що працюють в службі підтримки. Тому в якості ключових асоціацій, що будуть формувати образ продукту, обрано «масштабованість», «гнучкість», «автоматизація». Ключовими конкурентоспроможними позиціями продукту є:

а) адаптивність до середовища, в якому функціонує система – результати прийняття рішень в першу чергу залежать від користувачів, що є представниками певної віртуальної екосистеми;

б) легка масштабованість, що обмежується лише технічними можливостями;

в) рішення щодо контенту приймаються групою людей (в контексті прийняття рішень рішення, що приймається групою користувачів соцмережі, є оптимальнішим порівняно з рішенням, що приймається одним робітником служби підтримки) (табл. 4.17).

Таблиця 4.17 - Визначення стратегії позиціонування

<i>№ п/п</i>	<i>Вимоги до товару цільової аудиторії</i>	<i>Базова стратегія розвитку</i>	<i>Ключові конкурентоспроможні позиції власного стартаппроекту</i>	<i>Вибір асоціацій, які мають сформувати комплексну позицію власного проекту</i>
1	Гнучкість; надійність; масштабованість; низька вартість підтримки; здатність замінити групу фахівців, що працюють в службі підтримки	Стратегія диференціації	Адаптивність до середовища, легка масштабованість, групове прийняття рішень	Масштабованість, гнучкість, автоматизація

4.5 Розробка маркетингової програми стартап-проекту

Маркетингова концепція товару передбачає задоволення наступних потреб споживачів:

- а) оптимізація процесів;
- б) підвищення пропускну здатності;

- в) низька вартість підтримки;
- г) задоволеність користувачів контентом та процесом його контролю.

Для задоволення потреби компаній в оптимізації процесів продукт пропонує можливість делегування всіх процесів, що пов'язані з контролем якості контенту, системі консенсусу. Система в свою чергу оптимізує делеговані їй процеси. Ключовою перевагою системи консенсусу порівняно зі службою підтримки є її автоматизованість і відсутність необхідності залучення працівників. Для задоволення потреби компаній в підвищенні пропускну здатності продукт пропонує алгоритм, що забезпечує швидкий рух заявок в системі та мінімальний час їх розгляду. Ключовою перевагою системи консенсусу порівняно зі службою підтримки є її автоматизованість, що і забезпечує високу швидкість руху опрацювання заявок.

Для задоволення потреби компаній в низькій вартості реалізації контролю контенту продукт пропонує систему, що не потребує специфічних інструментів/ресурсів для функціонування. Для задоволення потреби компаній у високому рівні задоволеності їх користувачів продукт пропонує механізм залучення користувачів у процес прийняття рішень щодо прийнятності/неприйнятності контенту. Таким чином відбувається залучення користувачів мережі до процесу контролю того, що їх оточує в цій мережі і користувачі можуть власноруч контролювати середовище, в якому вони знаходяться. Ключовою перевагою системи консенсусу порівняно зі службою підтримки є той факт, що автоматизована система може збирати всеможливі показники та метрики напротязі усього процесу розгляду конкретної заявки і потім формувати багатоваріантну звітність, керуючись якою можна буде краще розуміти користувачів та їх потреби щодо контентного наповнення (табл. 4.18).

Таблиця 4.18 - Визначення ключових переваг концепції потенційного товару

№ n/n	Потреба	Вигода, яку пропонує товар	Ключові переваги перед конкурентами (існуючі або такі, що потрібно створити)
1	Оптимізація процесів	Автоматизація руху заявок в системі та процесу їх розгляду	Автоматизованість системи
2	Підвищення пропускної здатності	Висока швидкість руху заявок в системі	Автоматизованість системи
3	Низька вартість підтримки	Підтримка системи не потребує високих затрат навіть при масштабуванні	Автоматизованість системи
4	Задоволеність користувачів контентом	Залучення користувачів у процес прийняття рішень	Автоматизований збір відповідних метрик

Враховуючи той факт, що рівень цін на товари-замінники (тобто, кошти, що витрачаються компанією на організацію та оплату роботи працівників служби підтримки), як і рівень доходів цільових груп споживачів, може суттєво відрізнятись від компанії до компанії, оптимальним є встановлення фіксованої суми у розмірі 1 цкенту за одиницю опрацьованого контенту (табл. 4.19).

Таблиця 4.19 Визначення меж встановлення ціни

№ n/n	Рівень цін на товари заміни	Рівень цін на товари аналоги	Рівень доходів цільової групи споживачів	Верхня та нижня межі встановлення ціни на товар/послугу
1	Витрати на роботу служби підтримки індивідуальні для кожної компанії	-	Індивідуально для кожної компанії	1 цент за розгляд однієї скарги

В якості моделі збуту застосовується модель SaaS - модель поширення програм споживачам, при якій постачальник розробляє веб-програму, розміщує її й керує нею з метою використання її замовниками через інтернет. Тобто, компанії замовники платять за використання системи. Обсяги системи використання різняться від компанії до компанії (табл. 4.20).

Таблиця 4.20. Формування системи збуту

№ n/n	Специфіка закупівельної поведінки цільових клієнтів	Функції збуту, які має виконувати постачальник товару	Глибина каналу збуту	Оптимальна система збуту
1	- (модель SaaS)	- (модель SaaS)	- (модель SaaS)	- (модель SaaS)

Враховуючи обрану основу для позиціонування та визначену специфіку поведінки клієнтів, створено концепцію маркетингових комунікацій, що використовує в якості каналів розповсюдження тематичні ІТ-ресурси та соціальні мережі.

Для великих компаній, що є зацікавленими у нових технологіях, ключовими поняттями для позиціонування обрано автоматизацію, масштабованість та саморегулювання контенту. Концепцією рекламного звернення є примус до реакції на продукт (поширити інформацію про нього/дізнатись більше про продукт/спробувати продукт - для компаній).

Для середніх та великих компаній, що зацікавлені в оптимізації процесів в контексті контролю користувацького контенту, ключовими поняттями для позиціонування обрано автоматизацію, масштабованість, зниження витрат та підвищення пропускної здатності. Концепцією рекламного звернення є примус до реакції на продукт (поширити інформацію про нього/дізнатись більше про продукт/спробувати продукт - для компаній) (табл. 4.21).

Таблиця 4.21 - Концепція маркетингових комунікацій

<i>№ п/ п</i>	<i>Специфіка поведінки цільових клієнтів</i>	<i>Канали комунікацій, якими користуються цільові клієнти</i>	<i>Ключові позиції, обрані для позиціонування</i>	<i>Завдання реklamного повідомлення</i>	<i>Концепція реklamного звернення</i>
1	Відкритість до нових технологій	Соціальні мережі, тематичні ІТ-ресурси	Автоматизація, масштабованість, саморегулювання контенту	Розповісти про технологію та показати результати її застосування	Примусити реагувати на продукт
2	Зацікавленість в оптимізації	Соціальні мережі, тематичні ІТ-ресурси	Автоматизація, масштабованість, зниження витрат, підвищення пропускної здатності	Розповісти про технологію та показати переваги її застосування порівняно з існуючими рішеннями	Примусити реагувати на продукт

4.6 Висновок

В результаті проведеного аналізу було визначено, що проект з розробки системи консенсусу для децентралізованого контролю якості контенту має

можливість ринкової реалізації завдяки ряду факторів: унікальність продукту та можливість зайняти і розвивати власну нішу на ринку; потенційна можливість розвинути ринкову нішу до ринкового сектору; висока динаміка розвитку ринку; висока рентабельність ринку; наявність потенційного попиту з боку компаній-споживачів продукту. Проект має високі перспективи з огляду на те, що він не має прямих конкурентів, за рахунок чого порог входження низький, а бар'єри, які необхідно подолати, представлені наявністю продуктів-замінників (служб підтримки компаній). В якості альтернативи впровадження для ринкової реалізації проекту обрано комбінування двох альтернатив:

а) адаптація системи для можливості попереднього тестування та/або поступового впровадження компаніями-замовниками задля збільшення ступеню зацікавленості компаній-замовників та мінімізації їх ризиків;

б) залучення підтримки великих компаній.

Зважаючи на проведений аналіз ризиків та можливостей, залучення підтримки компаній паралельно з наданням цим компаніям тестових версій продукту для безпосередньої апробації ними системи є оптимальною стратегією для просування продукту і мінімізації ризиків, що пов'язані з недовірою компаній та самих користувачів до нової технології. Підсумовуючи результати проведеного аналізу, подальша імплементація проекту є доцільною.

ВИСНОВКИ

В дисертації розглянуті проблеми, що існують у сфері регулювання контенту в мережі Інтернет та запропонований підхід, що може бути застосований для вирішення даних проблем. Цей підхід покладено в основу створення децентралізованої системи оцінки якості контенту.

Запропоновано концепцію та розроблено математичну модель системи консенсусу для децентралізованого контролю якості контенту. Програмно реалізовано імітаційну модель системи консенсусу, за допомогою якої було проведено серію експериментів. Результати експериментів свідчать про працездатність запропонованої концепції консенсусу для децентралізованого контролю якості контенту.

Запропоновано варіант реалізації архітектури програмної платформи, що забезпечить горизонтальне масштабування системи консенсусу.

Запропонована система є простою, масштабованою і може бути використана в якості інструменту для здійснення саморегулювання користувацького контенту. Зокрема, вона може забезпечити:

- а) залучення спільноти користувачів до процесу оцінки контенту, що мінімізує суб'єктивізм та вплив людського фактору і зробить процес прийняття рішень про задоволення/відхилення скарг більш прозорим;
- б) створення глобального механізму регулювання користувацького контенту, що врівноважить та мінімізує вплив держав на процес фільтрації мережевого контенту та забезпечить гнучкість процесу його оцінки.

У майбутніх планах по розвитку цієї системи є врахування помилки суддями (наприклад, надати можливість змінювати рішення протягом деякого часу, і тільки після цього рішення суддів набуває сили), а також різна (мінлива) стратегія дій кожного учасника.

ПЕРЕЛІК ПОСИЛАНЬ

1. Network Enforcement Act (Netzdurchsetzungsgesetz, NetzDG) [Електронний ресурс]. – 2017. – Режим доступу до ресурсу: <https://germanlawarchive.iuscomp.org/?p=1245>.
2. Report on harmful communications and digital safety [Електронний ресурс]. – 2016. – Режим доступу до ресурсу: <https://www.lawreform.ie/news/report-on-harmful-communications-and-digital-safety.683.html>.
3. OSCE Representative on Freedom of the Media warns Germany social networks law could have disproportionate effect [Електронний ресурс]. – 2017. – Режим доступу до ресурсу: <https://www.osce.org/fom/347651>.
4. Internet Content Regulation in France and Germany: Regulatory Paths, Actor Constellations, and Policies [Електронний ресурс]. – Режим доступу до ресурсу: <https://www.tandfonline.com/doi/abs/10.1080/19331681.2013.803947>.
5. Міжнародний пакт про громадянські і політичні права [Електронний ресурс] – Режим доступу до ресурсу: https://zakon.rada.gov.ua/laws/show/995_043.
6. В Германии вступил в силу закон о крупных штрафах для соцсетей [Електронний ресурс]. – 2017. – Режим доступу до ресурсу: <https://roskomsvoboda.org/32565/>.
7. Руководящие принципы предпринимательской деятельности в аспекте прав человека [Електронний ресурс] – Режим доступу до ресурсу: https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_RU.pdf.
8. INTERNATIONAL FORUM "DEVELOPMENT OF PARLIAMENTARISM" [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <http://95.173.130.41/media/files/4pP0NhLGS6Z8RleOmWrFLT9d3MsFS0y.pdf>.
9. Звіт google про доступність сервісів і даних [Електронний ресурс] – Режим доступу до ресурсу: <https://transparencyreport.google.com/?hl=uk>.

10. Запити на видалення вмісту від державних установ [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <https://transparencyreport.google.com/government-removals/overview?hl=uk>.
11. Разработка прототипа конечного автомата для управления работой светофора на пешеходном перекрестке [Електронний ресурс]. – 2017. – Режим доступу до ресурсу: <https://kpfu.ru/portal/docs/F2117754092/diplom.Sisjov.pdf>.
12. Non-deterministic Finite Automaton [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: https://www.tutorialspoint.com/automata_theory/non_deterministic_finite_automaton.htm.
13. Деревья решений: общие принципы [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <https://loginom.ru/blog/decision-tree-p1>.
14. Деревья решений и алгоритмы их построения [Електронний ресурс] – Режим доступу до ресурсу: <http://datareview.info/article/derevya-resheniy-i-algoritmyi-ih-postroeniya/>.
15. Decision Trees - A simple way to visualize a decision [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: <https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb>.
16. Pattern: Saga [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <https://microservices.io/patterns/data/saga.html>.
17. Saga Pattern | How to implement business transactions using Microservices – Part I [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: <https://blog.couchbase.com/Saga-pattern-implement-business-transactions-using-microservices-part/>.
18. Реализация паттерна saga на C# [Електронний ресурс]. – 2018. – Режим доступу до ресурсу: <https://www.lshnk.me/2018/01/14/%D1%80%D0%B5%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F-%D0%BF%D0%B0%D1%82%D1%82%D0%B5%D1%80%D0%BD%D0%B0-saga-%D0%BD%D0%B0-c/>.

19. Microservices Using the Saga Pattern [Электронный ресурс]. – 2019. – Режим доступа до ресурсу: <https://dzone.com/articles/microservices-using-saga-pattern>.
20. A Guide to Consistent Hashing [Электронный ресурс] – Режим доступа до ресурсу: <https://www.toptal.com/big-data/consistent-hashing>.
21. System Design Interview Concepts – Consistent Hashing [Электронный ресурс]. – 2017. – Режим доступа до ресурсу: <https://www.acodersjourney.com/system-design-interview-consistent-hashing/>.
22. SQL vs NoSQL Difference [Электронный ресурс] – Режим доступа до ресурсу: <https://www.scylladb.com/resources/nosql-vs-sql/>.
23. Difference between SQL and NoSQL [Электронный ресурс] – Режим доступа до ресурсу: <https://www.geeksforgeeks.org/difference-between-sql-and-nosql/>.
24. WHEN TO USE SQL VS. NOSQL [Электронный ресурс]. – 2018. – Режим доступа до ресурсу: <https://www.integrant.com/when-to-use-sql-vs-nosql/>.
25. WHAT IS APACHE CASSANDRA™? [Электронный ресурс] – Режим доступа до ресурсу: <https://www.datastax.com/resources/definitions/what-is-apache-cassandra>.
26. What is Apache Cassandra? [Электронный ресурс] – Режим доступа до ресурсу: <https://www.scylladb.com/resources/introduction-to-apache-cassandra/>.
27. Is Apache Cassandra really the Database you need? [Электронный ресурс]. – 2018. – Режим доступа до ресурсу: <https://blog.knoldus.com/is-apache-cassandra-really-the-database-you-need/>.
28. Number of people using social media platforms [Электронный ресурс]. – 2019. – Режим доступа до ресурсу: <https://ourworldindata.org/grapher/users-by-social-media-platform>.
29. Social Media Statistics [Infographic] – Internet in Real Time [Электронный ресурс] – Режим доступа до ресурсу: <https://influencermarketinghub.com/social-media-statistics/>.

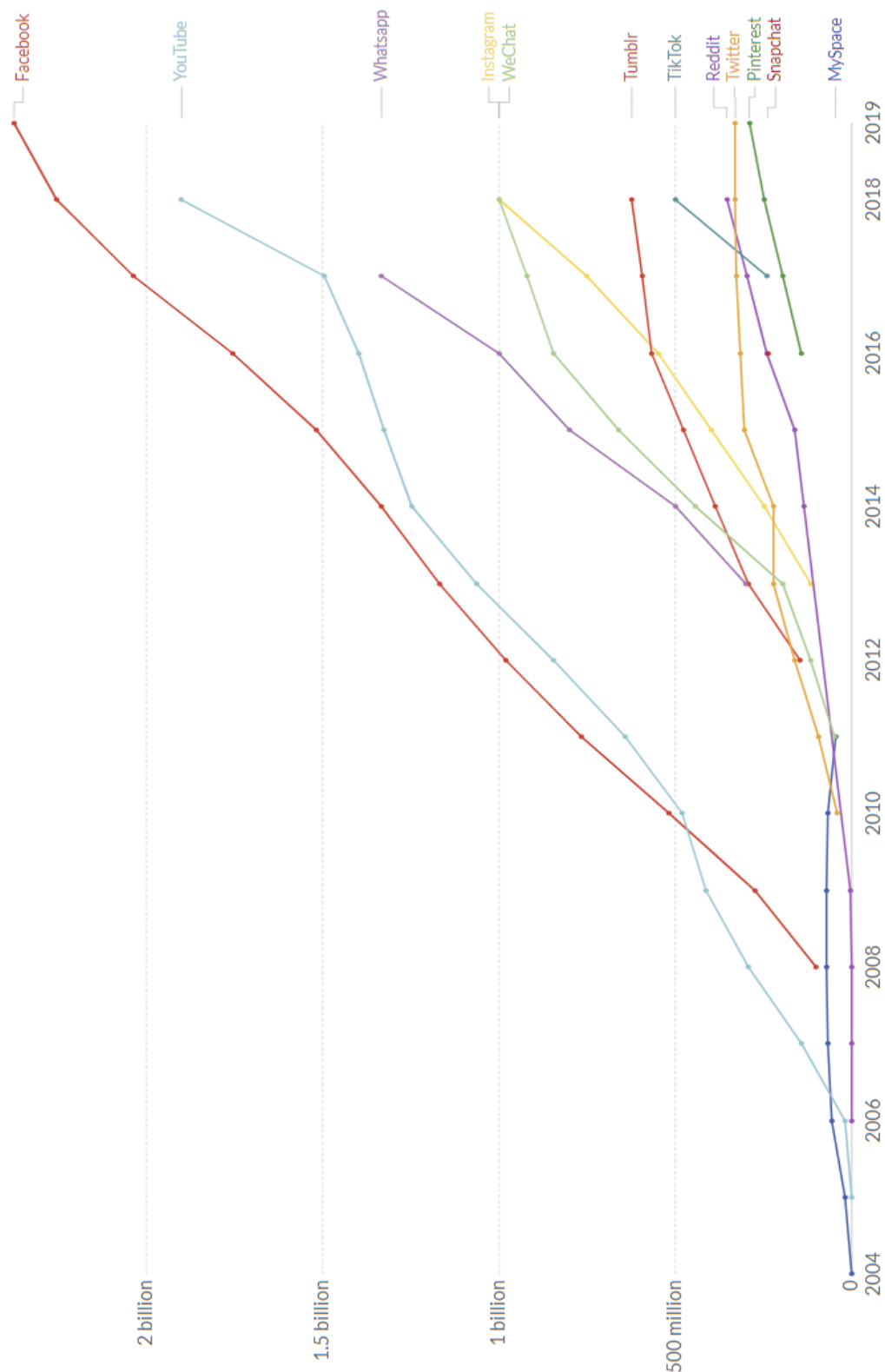
30. Most popular social networks worldwide as of October 2019, ranked by number of active users [Электронный ресурс]. – 2019. – Режим доступа до ресурсу: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.

31. Top 5 Social Media Predictions for 2019 [Электронный ресурс] – Режим доступа до ресурсу: <https://www.emarsys.com/resources/blog/top-5-social-media-predictions-2019/>.

32. Global Internet Report + Social Media Statistics & Facts 2019: How the Internet is Evolving [Электронный ресурс]. – 2019. – Режим доступа до ресурсу: <https://howsociable.com/blog/internet-and-social-media-stats-facts/>.

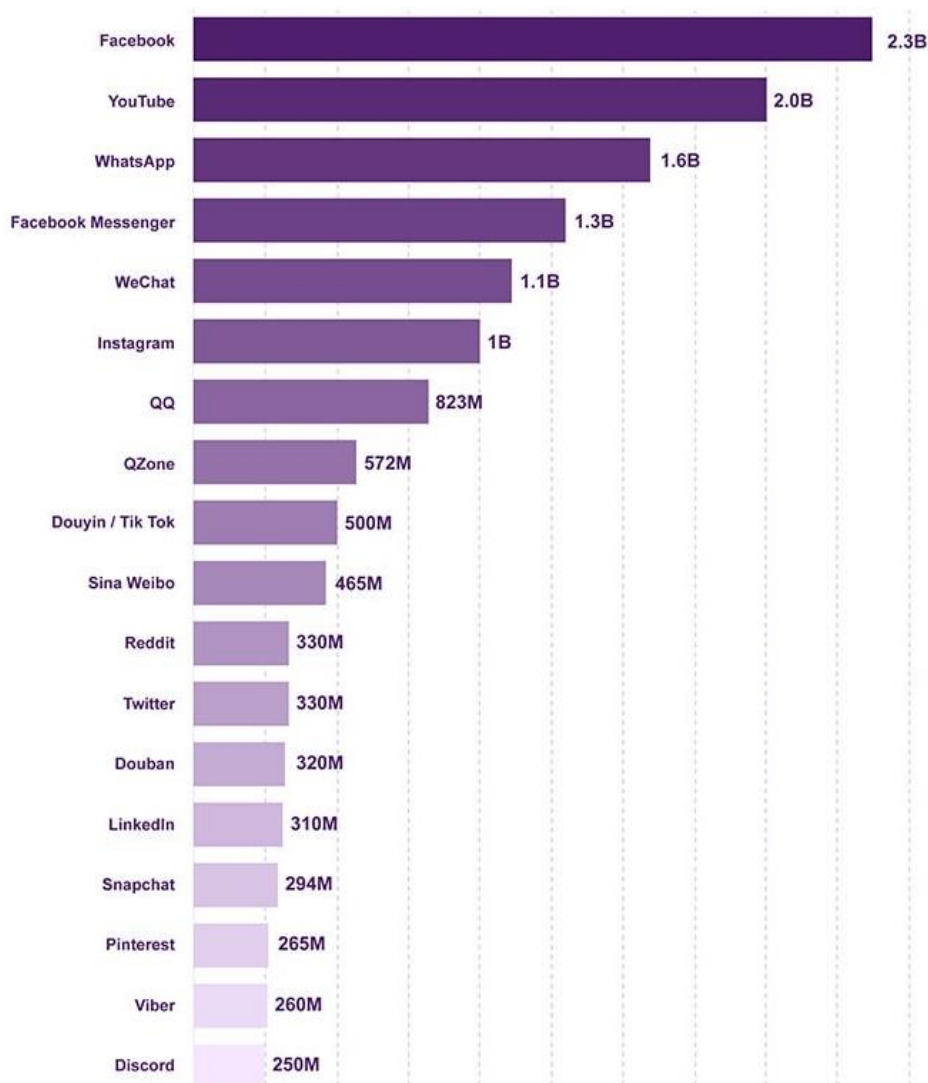
ДОДАТОК А ГРАФІЧНИЙ МАТЕРІАЛ

ПЛАКАТ 1 ДИНАМІКА ЗМІНИ КІЛЬКОСТІ АКТИВНИХ КОРИСТУВАЧІВ СОЦІАЛЬНИХ МЕДІА



Демонстраційний плакат до магістерської дисертації
«Моделювання системи для децентралізованого контролю якості контенту»
Магістрант Осипенко О.А.
Керівник Жданова О.Г.

ПЛАКАТ 2 КІЛЬКІСТЬ КОНТЕНТУ, ЩО ГЕНЕРУЄТЬСЯ КОРИСТУВАЧАМИ СОЦІАЛЬНИХ МЕДІА КОЖНІ 24 ГОДИНИ ТА ТОП СОЦІАЛЬНИХ МЕРЕЖ ПО СВІТУ, 2019

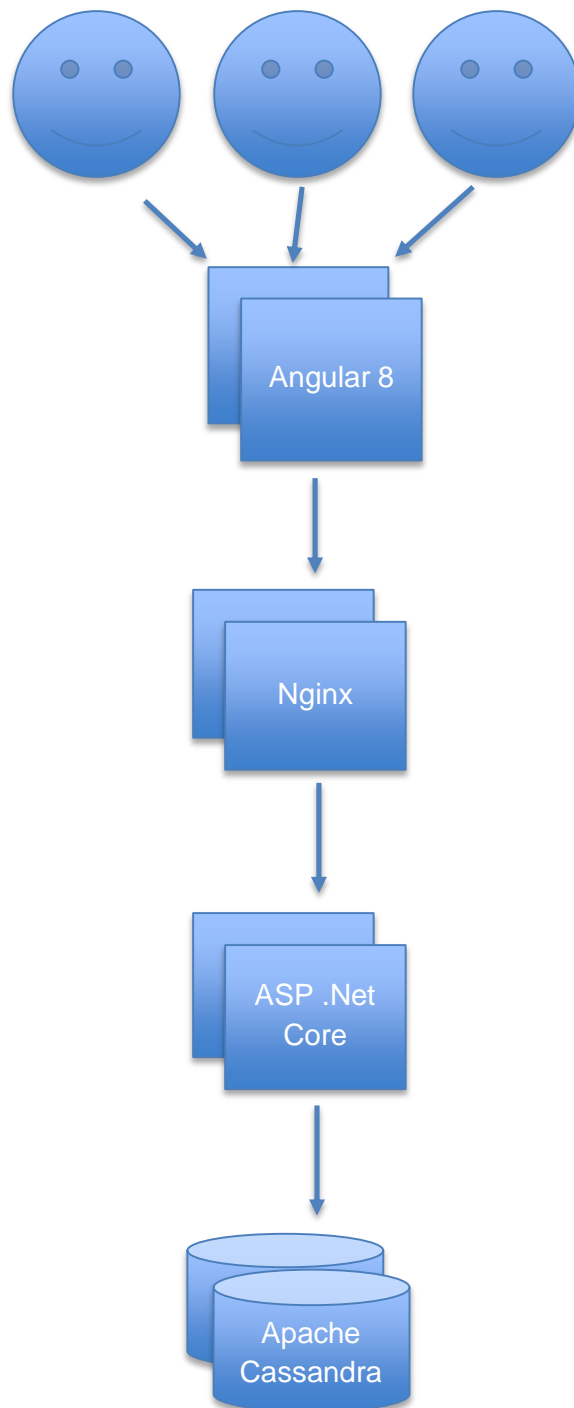


Демонстраційний плакат до магістерської дисертації
«Модельовання системи для децентралізованого контролю якості контенту»

Магістрант
Керівник

Осипенко О.А.
Жданова О.Г.

ПЛАКАТ 3 ДІАГРАМА АРХІТЕКТУРИ ПЛАТФОРМИ СИСТЕМИ КОНСЕНСУСУ

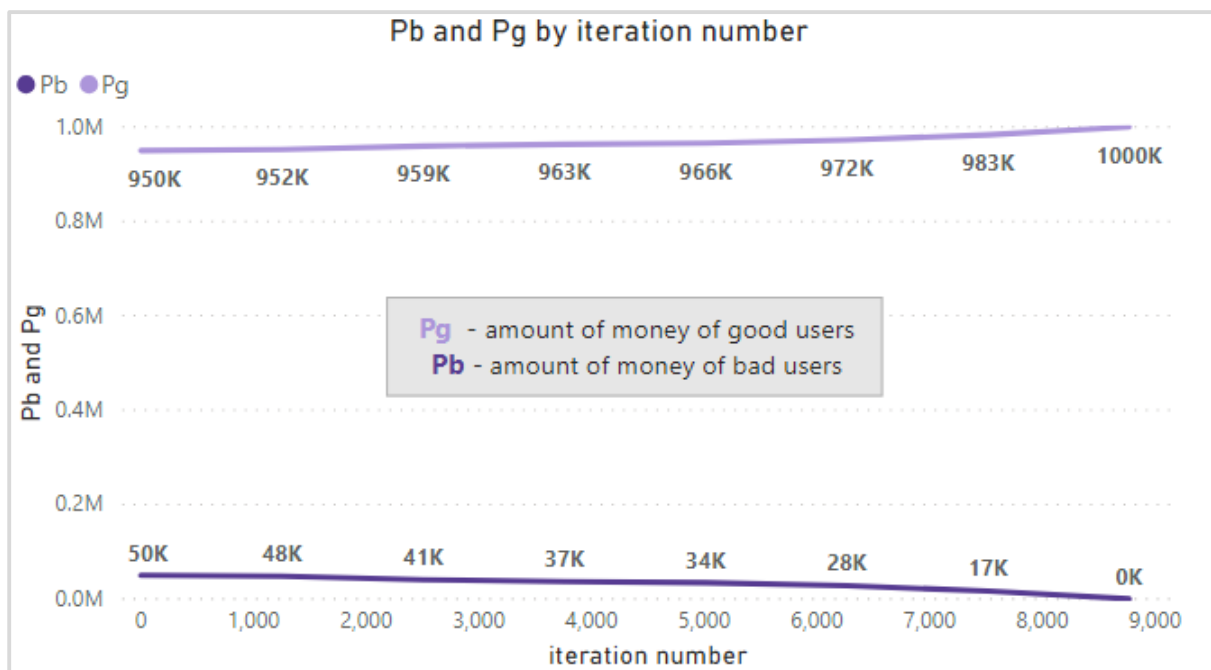


Демонстраційний плакат до магістерської дисертації
«Моделювання системи для децентралізованого контролю якості контенту»

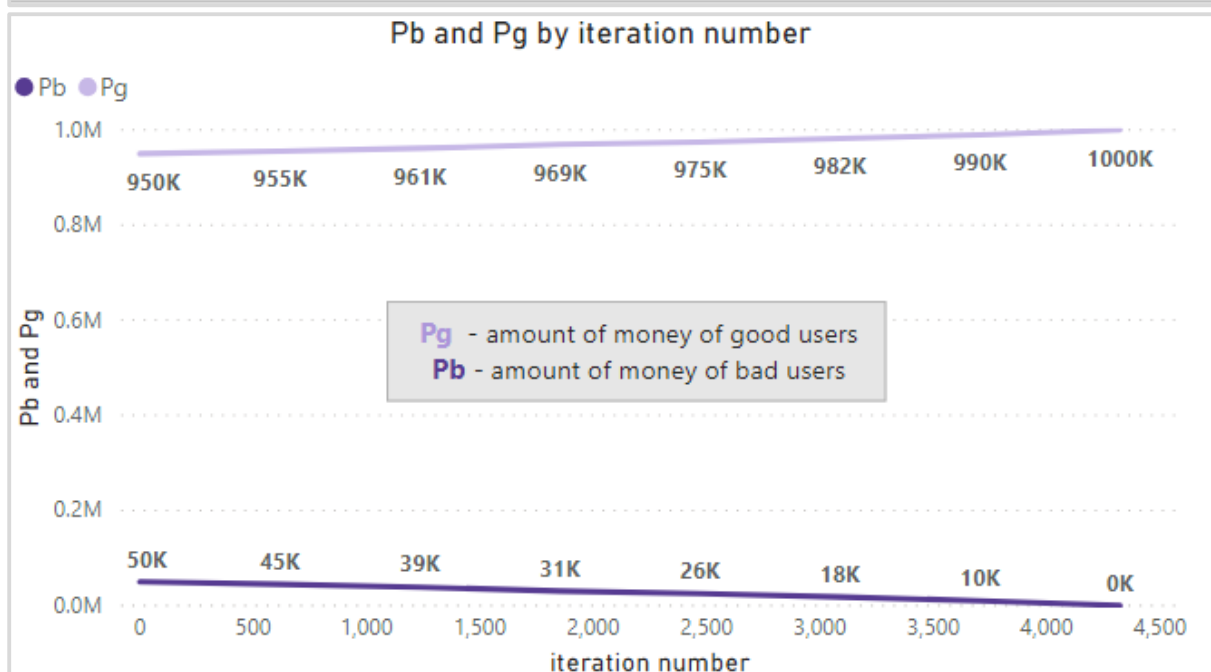
Магістрант
Керівник

Осипенко О.А.
Жданова О.Г.

ПЛАКАТ 4 ДИНАМІКА ЗМІНИ ОБ'ЄМУ ГРОШЕЙ КОРИСТУВАЧІВ В ЧАСІ ПРИ РІЗНІЙ КІЛЬКОСТІ СУДДІВ, 30% ПОГАНИХ КОРИСТУВАЧІВ ТА ВІДСУТНОСТІ АПЕЛЯЦІЙ



30% of bad users, no appeal, 3 judges



30% of bad users, no appeal, 7 judges

Демонстраційний плакат до магістерської дисертації
«Моделювання системи для децентралізованого контролю якості контенту»

Магістрант
Керівник

Осипенко О.А.
Жданова О.Г.